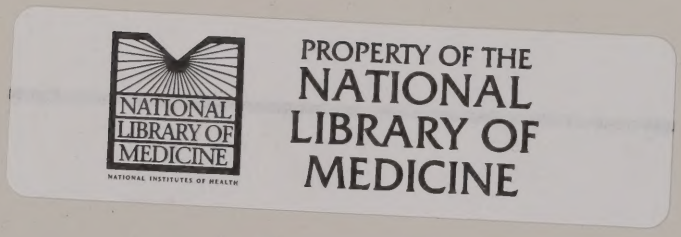
An Evidence-Based Health Care System:

The Case for Clinical Trials Registries

NIH Technology Assessment Workshop

December 6-7, 1993
Lister Hill Auditorium
National Institutes of Health



AN EVIDENCE-BASED HEALTH CARE SYSTEM: THE CASE FOR CLINICAL TRIALS REGISTRIES

OFFICE OF MEDICAL APPLICATIONS OF RESEARCH

NATIONAL INSTITUTES OF HEALTH

NATIONAL LIBRARY OF MEDICINE

LISTER HILL AUDITORIUM

BETHESDA, MARYLAND

DECEMBER 6-7, 1993

CONTENTS

AN EVIDENCE-BASED HEALTH CARE SYSTEM: THE CASE FOR CLINICAL TRIALS REGISTRIES

December 6-7, 1993

AGENDA

Monday, December 6

8:30 a.m.	Welcome and introductory remarks	W. Harlan D. Lindberg J. Clinton V. Pinn
9:00 a.m.	Background and goals of the workshop	J. Ferguson
9:10 a.m.	Definitions of RCTs and other experimental and observational studies	L. Friedman K. Dickersin
	Definitions of RCTs and other studies for indexing at NLM	P. Schuyler
	Search strategies for finding RCTs in MEDLINE	C. Lefebvre
10:20 a.m.	Discussion	
10:50 a.m.	The quality of levels of evidence for health care decisions	C. Hudgings G. Guyatt
11:20 a.m.	Discussion	
11:35 a.m.	Roles and uses of clinical trials registries—medical decisions, research, national and international cooperation	T. Chalmers I. Chalmers K. Dickersin
	Health care reform	W. Harlan
1:30 p.m.	Women's health and Congress	J. LaRosa
	Research integrity	E. Huth
	Experience outside U.S.	J.P. Boissel D. Moher J. Simes J. Ennis
	Patients' views	A.C. Baker P. Barr T. McCauley

Monday, December 6 (continued)

3:10 p.m.	Discussion	
3:40 p.m.	Examples of U.S. registries	
	• PDQ	A. Thurn
	• AIDSTRIAL/AIDSDRUGS	G. Dutcher & D. Katz
	• OMAR Registry	J. Ferguson
	• VA Registry	Y.K. Chan
4:35 p.m.	Core content of registries	K. Dickersin J.P. Boissel L. Colaianni
5:05 p.m.	Discussion	
5:30 p.m.	Adjourn	

Tuesday, December 7

8:30 a.m.	Registry maintenance, quality control	L. Colaianni J. James J. Porter P. Fayers J.P. Boissel
9:20 a.m.	Discussion	
9:40 a.m.	Issues, concerns, and cautions re establishing an NIH registry	R. Temple L. Edwards H. Schoolman J. LaRosa J. Porter
10:45 a.m.	Discussion	
11:00 a.m.	Summary	
12:00 noon	Adjourn	

WELCOME AND INTRODUCTORY REMARKS

DR. FERGUSON: Good morning. I am John Ferguson of the Office of Medical Applications of Research, the moderator of this conference. I would like to introduce Dr. William Harlan, Associate Director for Disease Prevention at the NIH and also Director of the Women's Health Initiative.

DR. HARLAN: Welcome to the meeting on the case for clinical trials registries as a part of the evidence-based health care system. Let me briefly state our goals for the meeting.

First, this is an open meeting that lends itself to discussion, questioning, and information sharing. We want to take a serious look at the case for a large clinical trials registry, to determine what it should contain, how it might be funded, how it can be accessed and used, and how it may relate to areas such as development of evidence for treatment, health care reform, and for the maintenance and monitoring of large clinical trials.

DR. FERGUSON: Our next co-sponsor is Dr. Donald Lindberg, Director of the National Library of Medicine.

DR. LINDBERG: I would like to call your attention to developments in the world outside of biomedicine, and particularly the world of high-performance computers and communications. The Internet now includes some 12,700 networks—not computers or persons—but networks of computers and people, some encompassing many states. There are over one and one-half million computers and 10 million people, a hundred countries, a thousand universities. This is a huge system and it is available to and supportive of the ideas that you want to bring into being. However, we tend to think about things in a biomedical context, and all that I described is not a biomedical context; it is all of science, of which we are a relatively small part.

Another interesting aspect that arises when comparing the work of biomedicine with that of science in general is that we are rightly critical (and that is the reason for this meeting) of the way biomedicine handles its affairs. For example, how well do we index things? make them retrievable? title them? Of all fields, however, biomedicine is perhaps alone in having discipline. Remember, too, that those large numbers of things I noted as being on Internet represent but 1 percent of the population. It is not designed as an elitist system, but it is, and the national information infrastructure that the President talks about is really aimed at the 99 percent not already served by these advanced systems. What you are meeting about is right and proper, and no doubt you will have much work to do after you reach a conclusion here.

In January 1991, a similar meeting was held to address the question of reporting results of clinical trials (particularly those sponsored by NIH). There was evidence that patients had been hurt

and their treatment gone awry or missed because clinical trial results were not given in an immediately effective fashion.

Dr. Ferguson constructed an interesting set of five or six case studies in which the trials were well done and, because of the data monitoring boards, were stopped as soon as the results were statistically valid. The trial results were then announced by various means, usually including press conferences, mailings, and the like. In every case, however, it appeared that whatever NIH had done was not sufficient, and there were criticisms. The problem was resolved like this.

The editors of the *Annals of Internal Medicine* (Edward J. Huth), *JAMA* (George Lundberg), and *New England Journal of Medicine* (Arnold Relman), all participated in that meeting. They agreed that, in cases where the results were important in the view of the director of an NIH institute (as the minimum case, not the only case) and where public health was threatened, the results should be announced in the fullest form necessary to make them clinically useful to doctors (not newspeople). Further, that journals (at least those represented by that panel) would not prejudice acceptance, but would submit such articles to their ordinary—in fact, even accelerated—peer review. That was an excellent result.

At the time of that meeting, audience opinions varied as to whether the circumstance I described—the urgent stopping of a clinical trial—would be a rare circumstance or a common one subject to gross abuse. In truth, there have been nine so-called "clinical alerts" since that time. Now when a clinical alert is announced for an NIH-sponsored trial (actually only 27 percent of the clinical trials in this country, although they tend to be the bigger and more expensive ones), there will be a press conference and a mailing, but NLM will immediately place the text in a clinical alert file. This informs anyone conducting a MEDLINE search (at least directly with us) that there has been a clinical alert by way of a "headline" announcing the results.

By itself, this information is insufficient for clinicians' use in treating patients. Two further levels of detail are, therefore, available. Online, a four- to 10-page abstract-like report is available, followed by an 800 number and additional sources of further information. NLM also makes the full text of this abstract immediately available on Internet. We fax it to 150 major medical libraries and send a mailing to the 4,000 members of the National Network of Libraries of Medicine, who often disseminate it further at the local level.

None of this means anything unless the work is valid and of high quality, and the titling and indexing are proper. NLM pays attention to these matters, as do the authors, editors, and investigators. Hence, my belief that your meeting will be both worthwhile and profitable, and again I welcome you and give you my best wishes for a good outcome.

DR. FERGUSON: I would now like to introduce Dr. Jarrett Clinton, the Administrator of the Agency for Health Care Policy and Research.

DR. CLINTON: Let me add my welcome to our colleagues from NIH who have been quite good in incorporating us into this activity.

The Agency for Health Care Policy and Research is enjoying its fourth year of authorization by Congress, having been created to address apparent variations in practice in the United States and elsewhere. Similar patients were being treated differently, with little attention paid to their short- or long-term outcomes. At times even basic morbidity and mortality information was not captured adequately.

At issue was what might be done to build on the science base of the practice of medicine. The focus was not so much on the efficacy trials, randomized clinical trials, and knowing what works best in controlled environments, but rather on the generalizability of what works with real patients, with all their co-morbidities, and with real physicians—not all of whom may have the same degree of information, nor make the same effort to stay up to date with new information.

Our program takes two formats, with one being the research effort. Key among our medical effectiveness program research is what we call "patient outcome research teams." These are large, multidisciplinary, generally multi-institutional efforts to examine common diseases in the United States, and include back pain, myocardial infarction, cataracts, and benign prostatic hyperplasia among the 14 that are currently underway. The goal is to review the literature and databases (including claims and medical records databases) to determine what works best and transform it into the clinical practice of America.

We have been astounded by how much we have learned, and by how much we have learned that we do not know. For example, randomized clinical trials, particularly in the cardiovascular area, have produced enormous amounts of information. There has been a limit to what we have been able to contribute to cardiovascular knowledge in areas other than diagnostic imaging, where there is still considerable variation in practice across the country. We are unsure of the outcomes or implications of this, other than cost.

In another example, the prostate group working at Dartmouth learned that many clinicians, urologists, and patients are quite comfortable with radical prostatectomies. The radical prostatectomy rate in the United States has increased strikingly in the last several years, with no good evidence of its effectiveness. It has been stimulated by tremendous interest in the management of prostate cancer.

The fact remains, however, that there are no randomized clinical trials on prostate cancer. (One was recently initiated here at the National Cancer Institute.) So we have much to learn, even in areas we thought were well understood a few years ago. Our investigators need the information from

randomized clinical trials t[illegible]plement the information from other databases, and to provide t[illegible] average practitioner with d[illegible]n what seems to work best for the American public.

The second phase of our activity focuses on developing clinical practice guidelines. It is not research but, rather, the synthesis of what we know. Again, it is aim[illegible] at practitioners. The goal is to combine information availa[illegible] in written and electronic formats from databases and professional judgement. In collaboration with the National Library of Medicine, we hope to access existing information and transmit it to the practicing world.

We have been enormously surprised, both by the magnitude of literature that is available on a given subject, and by how little it contains. This has compounded the problem of combining the general sense of this research effort into something that is practical and generalizable for practitioner and patients alike. Therefore, we welcome an opportunity to understand more completely how we can develop a framework for collecting information from randomized clinical trials, both completed and ongoing. We look forward to your recommendations in these areas.

DR. FERGUSON: It is now my pleasure to introduce Dr. Vivian Pinn, the Director [illegible]e Office of Research on Women's Health.

DR. PINN: I am delighted that the Office of Research on Women's Health was able to join in sponsoring this workshop and add my welcome to those who have preceded me.

As you are aware, our office was established in 1990 with the purpose of developing information related to gaps in knowledge about women's health, to ensure the participation of women in clinical trials conducted by the NIH, and to help advance careers of women in biomedicine. One of the important expectations of our office is that information generated about women's health from our investigations be made available, not only to the scientific and health care professional community, but also to women and those who are interested in women's health.

Congress has been particularly persistent in assuring that women's health becomes an integral part of the biomedical agenda here at the National Institutes of Health. Much of the language related to women's health research in the 1993 Revitalization Act of NIH involves inclusion of women in clinical research, but there is one particular area I would like to bring to your attention.

The act directs that the Director of the NIH, in consultation with the Director of the Office of Research on Women's Health and the Director of the National Library of Medicine, establish a data system including a registry of clinical trials of experimental treatments for the collection, storage, analysis, retrieval, and dissemination of information regarding research on women's health. As part of that data system, we are also charged with establishing a registry of clinical trials, which must include information on sex, age, ethnicity or race, and the location of the trial site or sites. The language further directs the NIH Director, again in consultation with the Office of Research on Women's Health

and the National Library of Medicine, to establish, maintain, and operate a program to provide information on research and prevention activities of the national research institutes that relate to women's health.

Therefore, the efforts of this workshop will be particularly important to us in addressing our Congressional mandate. Further, the development of the clinical trials registry will be especially relevant and important as we consider the 1993 guidelines for the inclusion of women and minorities in clinical research. These new guidelines are due to be released within the next few days.

The Revitalization Act of 1993 also directed us to change the NIH policy which has existed since 1990 for the inclusion of women and minorities in clinical research. These guidelines will address the issues of the inclusion of not only women and minorities, but also of subgroups of minority populations in Phase III clinical trials, so that a valid analysis of differences can be accomplished. The guidelines now further state that cost cannot be considered a factor in the exclusion of women and minorities as participants in clinical trials, and also require the NIH to support outreach for the recruitment and retention of women and minorities into such studies.

These guidelines will be effective for all funding beginning in FY 1995, meaning that grant applications coming in as early as February 1994 must comply. If a clinical trials registry were available today, it would assist clinical investigators in knowing what studies have been done, whether or not women were included, and which populations of minorities were included, so they would be in a better position to design their studies to meet these new guidelines.

BACKGROUND AND GOALS OF THE WORKSHOP

John Ferguson

Office of Medical Application of Research

DR. FERGUSON: I would like to provide some background for this workshop. There have been a number of meetings of a group from the Society for Clinical Trials on the subject of registries of clinical trials, and I was asked to give a presentation to them a couple of years ago. It is interesting to note that the first president of the Clinical Trials Society was Dr. Harold Roth, recently retired from NIDDK, and a recent past president was Dr. William Friedewald, who was instrumental in my coming to NIH. Thus, there has been great interest from the outside community on this.

OMAR has had a registry at NIH for a number of years (which will be discussed later), and there have been difficulties with its use and maintenance. NIH has other sources for information on ongoing clinical trials, and there are also the MEDLINE, PDQ, and AIDSTRIALS/AIDSDRUGS databases. More information will be provided on these later, and we have just heard about the NIH Revitalization Bill and its mandate for women's health.

The purpose of this workshop is information gathering. We are looking for how-to information. We will review the definitions of clinical trials and varieties thereof, along with the quality and levels of evidence for health care decisions. I think this aspect is vital, and Dr. Clinton has spoken a bit about it. We will gather information on registries of ongoing trials (that is, those that are not yet published), and hear about how to retrieve information from trials that are published.

We will explore NIH's role in establishing these registries, along with some of the concerns and needs of the health community in general, which is all of us. For example, if something is in a registry, would an insurance carrier not reimburse for it because it is part of an ongoing trial and therefore an experimental or investigational procedure?

We will also consider options for a possible action plan. What can, could, and should we do here at the NIH?

There are several reasons for having trials registries, and we will hear more about them during the course of this meeting. One is that they can help prevent needless duplication. They also can help promote necessary replication. Perhaps, for example, many trials have been completed but the information for any individual patient is still incomplete. Replication is sometimes necessary to combine results of trials. Registries can also help in minimizing publication bias and in performing systematic reviews of the literature.

There are also ethical reasons for maintaining clinical trials registries. They can be helpful in patient accrual by providing patients with information on ongoing trials and how to enroll in them.

Registries can be used as a basis for methodologic research, cross-fertilization among trialists, and in planning reasonable health policy in our evidence-based health care system.

We will spend most of our time today on registries of unpublished or ongoing trials. For the purposes of this workshop, I will emphasize that what goes into these registries is not the trial results. That may be a possibility, but registries in general should contain trial characteristics; that is, when it started, how many patients, the type of trial, and so on.

You will also hear about published trials. It is difficult to get them all from MEDLINE, and we will explore some of the reasons for that. Is it possible to retrospectively tag things in MEDLINE? Is it useful to have published clinical trials in the registry from before 1966 when the electronic MEDLINE was established? To what extent should we know about trials that were published in other journals and not indexed by MEDLINE or the National Library of Medicine?

I would like to now turn this over to Dr. Larry Friedman from the National Heart, Lung and Blood Institute.

DEFINITIONS OF RCTs AND OTHER EXPERIMENTAL AND OBSERVATIONAL STUDIES I

Larry Friedman

National Heart, Lung, and Blood Institute

DR. FRIEDMAN: I have been asked to talk about definitions of clinical trials. As background, I will review different kinds of studies from broad to narrow. The broadest are studies in human beings or human-based biologic materials. This is not a rational definition of a clinical trial. The second definition is narrower, describing studies in living people. Unfortunately, this encompasses observational studies—for example, cohort studies, case control studies, and case reports, not just intervention studies. Again, I think this is too broad. For those who are unclear, a cohort study or an observational study is a longitudinal study where people are naturally exposed or not exposed to various risk factors, and the outcome is observed. A case control study is a retrospective study in which patients with a disease or condition are evaluated for possible exposures that might have occurred previously.

The third definition, intervention studies, is still narrower but may include studies without comparison groups. I maintain that this definition is still too broad.

The fourth definition, comparison studies, is getting close. However, this may also include historical controls. I am not sure that we want to include them in a definition of clinical trials.

The fifth definition is concurrent controls where the control and intervention groups are observed at the same time. These are not necessarily prospective studies, although we may not want to include that in our definition of a clinical trial.

The sixth definition, prospective control studies, is getting closer. The question now becomes the selection of the control group. Must it be randomized, or can it be selected in other ways, perhaps matched in some fashion? The seventh definition is randomized prospective studies. Randomization must use proper procedures, not selecting the day of the week, birthdate, or other easily manipulated factor. This may include a variety of early studies and studies of small numbers of patients. As described below, this definition includes what most people consider to be clinical trials.

Finally, I use "Phase III" as a shorthand for large clinical trials designed to answer a clinical question in a clear and semi-definitive fashion. Note that any of these areas might include crossover and factorial design studies as well as classic parallel studies.

The implications for these definitions are broad. Ranges would encompass more studies, including many whose validity we would question. The narrow definitions will include only the better studies and

those that are more easily found. However, some key studies may be missed. There is a trade-off between sensitivity and specificity.

I will review existing published definitions of clinical trials. They are all similar, and most include studies of prospective, randomized, and "Phase III" studies.

The first example is one that Curt Furberg, David DeMetz and I published some years ago. We defined a clinical trial as a prospective study comparing the effect and value of interventions against a control in human subjects. The points we made were that the study must be prospective, controlled, of human beings, and an active intervention.

Curt Meinert has defined a clinical trial as a planned experiment designed to assess the efficacy of a treatment in man by comparing the outcomes in a group of patients treated with a test treatment with those observed in a comparable group of patients receiving a control treatment; where patients in both groups are enrolled, treated, and followed over the same time period. The words are different, but I think the context is the same as in the previous definition.

Stewart Pocock has defined a clinical trial as any form of planned experiment which involves patients and is designed to elucidate the most appropriate treatment of future patients with a given medical condition. This is similar to the preceding definitions.

Bulpitt has rewritten Bradford Hill's definition. He defined a clinical trial as a carefully and ethically designed experiment which includes the provision of adequate and appropriate controls by a process of randomization so that precisely framed questions can be answered. He added two points to the previously definition. One, he included randomization as part of the definition and, two, he specified that the trial be ethically designed. I hope that the other trials would also have been ethically designed, even though this was not specified in the definition.

Other definitions have been published. Alvin Feinstein, in *Clinical Epidemiology*, defines clinical trials as "forward-directed studies in which the people under observation are followed as a cohort from baseline state through imposition of maneuvers to the incurrence of outcomes. This type of cohort research is called a clinical trial when the maneuvers are imposed with an experimental plan." He contrasts this with an observational cohort, where the maneuvers are imposed under ordinary circumstances.

Finally, the *Dictionary of Epidemiology*, second edition, 1988, defines a clinical trial as "a research activity that involves the administration of a test regimen to humans to evaluate its efficacy and safety." This continues by stating that the term is subject to wide variation in usage, from the first use in humans without any control treatment, to a rigorously designed and executed experiment. I think that, until recently, there has been a problem in that the term clinical trial was been applied to the kinds of situations

presented at the beginning of this talk which most people today would not consider to be a true clinical trial.

My final recommendation is that a clinical trial should have the following characteristics: It must be prospective, controlled, and an intervention which is administered in a standard fashion. Also, for the purposes of a registry, we ought to consider using randomization as a key characteristic.

DEFINITIONS OF RCTs AND OTHER EXPERIMENTAL AND OBSERVATIONAL STUDIES II

Kay Dickersin
University of Maryland

DR. DICKERSIN: To begin, I will describe why a definition matters. Most of what I will say is in the context of a registry of clinical trials, rather than why clinical trials are important. First, we need to agree whether we mean published or ongoing trials. Many of my remarks are aimed at the registry of published trials, although they can be applied to ongoing trials as well.

The existing mechanisms for identifying trials, such as MEDLINE and other databases, use various definitions. In addition to defining what we are looking for, we must incorporate these different definitions to identify existing trials that meet our internal criteria. Finally, to develop a registry or registries, we need to set eligibility criteria. We therefore need to make a decision about what we consider a clinical trial to be.

I think the criteria should include the following characteristics which follow nicely from Dr. Friedman's comments. First, a clinical trial is a comparison of at least two forms of health care, possibly including a placebo or nothing unusual. An intervention or no intervention would also be two forms of health care. Next, there must be concurrent study groups. The study must have been planned and not result from analyzing historical records. I think an unbiased method of treatment assignment, random allocation, also must be used. This helps eliminate selection bias and allows researchers to compare two forms of treatment in two types of patients, whose treatments are based on some factor related to their prognosis.

Curt Meinert defined a clinical trial in 1991 as "any therapeutic, diagnostic or preventive study involving humans, which compares concurrently one intervention (a drug, device or procedure) to another intervention, placebo or no intervention, in order to determine their relative safety and efficacy." This definition of a clinical trial does not mention randomization.

Curt Meinert's definition of random allocation is the process of assigning persons to a group, using a random process such as a table of random numbers. This is not a haphazard process; it is a distinct process that meets a statistical definition of random.

As mentioned, the existing means of finding trials may use different definitions of a clinical trial. Here is a list of some of the databases that are currently in use.

MEDLINE is perhaps the most familiar database because it is easy to use and we tend to identify trials from the journals in which we search. In addition, there are other MEDLARS databases that could be used to identify trials. EmBASE, by Elsevier, is the European counterpart to MEDLINE. It

uses a completely different vocabulary. I am not sure they have had a term for clinical trial. Other non-American bibliographic databases may or may not index under terms such as clinical trial or randomized clinical trial. CRISP, the database of all Public Health Service funded studies, has its own definition of a clinical trial. There also are various registries of research, including registries of trials or other types of research and bibliographies of review articles. These do not index clinical trials but may be used to find them, although, again, the author's definition of a clinical trial may be different from ours. Finally, clinical practice guidelines use clinical trials and other types of studies as a basis for their conclusions. This may be another source for randomized clinical trials.

If we are going to establish a registry of published trials, we need to set criteria for choosing clinical trials. The first criteria is that our definition must be easy to understand. If it can not be understood, it can not be applied. The second criteria is that the definition must be easy to apply. Even if a definition is understood, it cannot always be applied to a given trial in the literature. This is the issue of sensitivity versus specificity. I would argue in favor of being sensitive and including studies that may not meet a rigorous definition of a randomized trial. This would allow users to know that they have found all possible clinical trials and to decide for themselves which trials suit their purposes.

The Cochrane Collaboration, which will be presented later, is involved in registering both published and unpublished trials. The criteria are that individuals, or other units such as schools or communities, followed in the trial must have been definitely or possibly assigned prospectively to one or more alternative forms of health care, using either random allocation or quasi-random allocation such as alternation, date of birth, or hospital record number. This registry includes quasi-randomized studies and studies that may have used questionable randomization procedures in order to avoid missing possible trials.

There are two main problems with applying these definitions. First, a paper that describes a trial under consideration for inclusion in a trials registry may lack explicit description of the methods that were used. Furthermore, the explicit description in the article may not be correct. For example, the author may state that the trial was randomized when, in fact, it may not have been.

I am going to present some data from a small study that we did with Roberta Scherer, Peri Schuyler, and co-workers to form a gold standard of randomized trials in the vision literature, ophthalmology and optometry. MEDLINE search and hand searching of 66 journals was used to attempt to identify all randomized trials in that field for 1988. Through these searches, we found 1,513 reports that were possibly randomized trials. By reviewing them we determined that 1,294 definitely were not randomized trials; 201 definitely were randomized trials; and 18 that we could not

classify from the paper, but were determined to be randomized trials after we confirmed the study design by letter to the author.

Of the 219 confirmed vision trials identified from the literature, we identified 189 using MEDLINE. Thirty were identified purely by a hand search. We investigated why they were they not identified using MEDLINE.

In 1988, there was a MeSH tag clinical trial. We explored the likelihood of finding randomized trials (what we called randomized trials) under this tag. Of the 189 trials found using MEDLINE, 77 had the tag clinical trial. In all, 79 had the tag clinical trial, and two of them, for some reason, were not found using MEDLINE. Therefore, everything we called a trial was not tagged clinical trial in MEDLINE. This example demonstrates the difficulty of recognizing trials while indexing at the National Library of Medicine (NLM); one cannot always tag trials the way they should be tagged.

We reviewed the 189 studies to determine what information was provided to identify them as clinical trials. Of those with the clinical trial tag, the majority had words or terms in the title and abstract that identified them as clinical trials such as "randomized clinical trial." A number also had identifying terms in the methods and the body of the text. One study had no descriptive terms, but somehow the indexer knew it was a clinical trial. Three studies have not yet been translated, so we do not know where that information is. However, we know from a letter to the authors that they were randomized clinical trials.

For those not tagged by NLM as clinical trials, the information was most often in the methods section. Sometimes these should have been tagged, because it was quite clear that they were randomized trials. The information was not in the title or abstract, and presumably the indexer did not read as far as the methods. Sometimes the information in the methods and elsewhere was not very clear, and for this reason, they were unable to be tagged.

Some of the descriptors used in the methods sections of untagged clinical trials include "assigned randomly", "computer generated randomization", or "randomized". We felt these terms were clear, and in those cases, they should have been indexed. When the author said "the patients were randomly chosen", "chose at random", or "patients were divided at random", it is not clear that these are randomized trials. They might have said "alternate patients allocated" or "alternately", and if the definition of randomized trial does not include quasi-randomization, these would not be identified. One of the problems was that in some eye trials, the eye that received the treatment was often randomized. So a definition of randomized clinical trial would have to include the possibility that eyes, rather than people, were randomized. At other times an article contained terms such as double-blind, double-masked or cross-over, which some of us might think indicates a clinical trial, but the

indexers might not. The main point is that the authors need to be more careful in choosing terms to describe their study and methodology.

Next, I will present discouraging results from a different study. Not only do authors sometimes not describe what they have done explicitly enough to allow indexing, but sometimes they do not seem to know what they have done. In this study by Roberta Schuyler and others at the University of Maryland, we took randomized clinical trials in the vision field and identified abstracts presented in 1988 and 1989 at ophthalmology meetings. We chose these abstracts because the authors described their studies as randomized clinical trials, a clinical trial, or perhaps they did not use either term, but they said something that made us think it was possibly a randomized trial. We then wrote to all the authors of the abstracts and asked them if their study was indeed randomized. If it was, we asked what method of randomization was used, to ensure they knew what randomization was. The sad thing is that, for six of the studies where the authors had explicitly said it was a randomized clinical trial, they responded to us in our questionnaire that the study was not randomized. This is an example, then, not only of failing to be explicit, but of being explicit and not being correct. Similarly, in five cases where the authors had not said anything to identify their study as a randomized trial but we thought it might have been, they responded that it was a randomized trial.

This discussion of study definitions presents a starting point for us in perhaps one of the most difficult areas in establishing a register of trials.

DEFINITIONS OF RCTs AND OTHER STUDIES FOR INDEXING AT NLM

Peri Schuyler

National Library of Medicine

DR. SCHUYLER: Rather than focusing solely on the definitions, I have divided my presentation into four parts. First, a brief picture of the scope and coverage of the MEDLINE databases at the National Library of Medicine (NLM) will be presented. My purpose is to provide perspective with respect to the total number of published journals in biomedicine, to describe the growth of the database, especially in terms of clinical trial coverage, and to present a brief summary of the maintenance and control measures that we take to ensure the quality and reliability of data in the database.

Second, I will present the evolution of indexing policy on clinical trials from 1966 to the present and the effect of this evolution on the definition and interpretation of clinical trial data. Third, I will present an overview of the terminology available in MeSH, Medical Subject Headings, the Library's controlled vocabulary. And last, I will present a summary of the steps taken to increase the accessibility of clinical trial literature and our commitments for the future.

Indexing at the National Library of Medicine centers on the over 3,700 journals that are included in the MEDLINE database. These journals span a domain of biomedicine that has been expanding steadily since the mid-1970s to encompass today the delivery of health care, molecular biology, health services research and technology assessment, and space life sciences. As each of these areas has risen to prominence in the scientific and public eye, the complexion of MEDLINE and its back files has changed accordingly. In many cases additional, more specialized databases have been created to supply the coverage necessary to support research and practice in these fields.

There are three points I would like to make about NLM's coverage. First, the 3,700 titles that we presently index represent less than 20% of the total number of journals now being published in biomedicine. So, although these 3,700 titles yield approximately 350,000 citations each year, they do not represent the bulk of what is published.

Second, since we do not have the resources to index everything, nor is it clear that would be the best policy even if we had those resources, how do we determine what titles to include? NLM has always relied on outside consultants to advise us on journal selection. But beginning in 1990, this process was formalized with the establishment of an NIH-chartered committee called the Literature Selection Technical Review Committee. The Committee's members represent all aspects of the health sciences, from education and research, to clinical practice and librarianship. Every journal reviewed by the Committee is assigned both a primary and a secondary reviewer, and at least four of the most

recent issues are examined to assure that a reasonable sample of the journal's contents are represented. The Committee reviews close to 400 journals a year, but generally recommends fewer than 80 for inclusion in MEDLINE.

The third point regarding coverage is that we are committed to reviewing every journal. Even journals being indexed are reviewed. If there are available journals that represent clinical trial data, please let us know, and we will be certain to present them to the Committee for review. In 1966, when MEDLINE began, we indexed only about 2,600 titles, and they yielded between 100,000 and 120,000 citations each year. Today's 3,700 titles yield, as I mentioned before, about 350,000 citations. This means that not only are we indexing more journals, but that those journals are yielding about twice as many articles a year as they did in 1966—from about 50 a year then, to about 100 a year today.

This growth rate is also reflected in the clinical trials literature, which has increased from less than 1% of the database in 1966, to nearly 3% today. The total volume of clinical trial data in MEDLINE from 1966 to the present is around 120,000 citations. The current MEDLINE file, which covers the period from 1990 through the present, contains about 38,000 clinical trial citations.

An important thing to realize, however, is that our databases are fluid. We engage in constant maintenance activities, both on an annual, global basis, and on an individual record basis. The significance of this is that errors can be corrected and new information added. Original articles can be linked to commentary on those articles, and duplicate citations, retractions and published errata can be identified. Clearly the capability for linking clinical trials in registries with their published reports, or for adding key indexing terms to those citations, is in place.

Our indexing policy with respect to clinical trials has evolved over the past several years. In the early years of the database, our policy was quite strict. Strong emphasis was placed on control. Indexers were cautioned that, just because an author claimed a study to be a trial, it might not be a true controlled trial. Unless there was explicit description to that effect, they were not to designate a study as such in indexing. It is an important tenet in indexing policy, regardless of the subject area, that the indexer neither interpret nor anticipate the possible intent of the author. This has lead to high reliability and precision, but contributes to a possibility of under-identification of potentially eligible trials. As Iain Chalmers has said, the Library's control vocabulary has been applied with a concern for specificity, rather than being designed to maximize sensitivity. Indexing errors under these conditions are those of omission, rather than of commission. Another view is that you might not get everything that you want, but what you get you will probably like very much.

The continuing dialogue between NLM and various researchers has led to a greater understanding of the responsibilities on both sides and was clearly the impetus behind indexing

policy's move toward a broader interpretation of clinical studies in 1991. This is a different direction from Dr. Friedman's, but it is designed to allow researchers to determine the quality or eligibility according to their own needs and to maximize their ability to retrieve data.

Since 1991, the index section has issued a series of technical memoranda that describe the various elements of trial design and construction, emphasizing the pre-planned experimental nature of these studies and de-emphasizing the necessity for the description of controls. The current definition is a modification of Meinert's definition. As a result of this change, coupled with the expansion of clinical-trials-related vocabulary that I will describe briefly below, there has been a significant increase in the percentage of articles to which indexers have assigned a clinical trials heading. In the studies Kay Dickerson referred to, we had compared indexing for a subset of the randomized controlled trials that she and I had worked on previously. The subset in ophthalmology compared trials published in 1988 and trials published in 1991 when the trials had an explicit mention of randomization in the title or the abstract. In the subset of trials, the percentage identified by the indexers when a statement concerning randomization appeared in titles or abstracts increased from 69% in 1988 to 89% in 1991. I repeated the experiment last week for the purposes of this talk, and the improvement continues: 92% of the ophthalmology titles carrying a statement about randomization or random allocation in titles and abstracts were indexed to an appropriate clinical trial heading.

In describing these experiments, I have alluded to the presence of a statement of explicit randomization in titles or abstracts. Remembering that the job of indexers is to report, and not to interpret, it should be clear that the more straightforward and explicit authors are in characterizing their studies of clinical trials, the greater the likelihood that those studies will be assigned the appropriate clinical trial heading. We, for our part, shall continue to monitor indexing operations and to ensure that indexers are conversant with the characteristics and tools of clinical research.

The world's greatest indexers cannot index in the controlled vocabulary environment without the headings necessary to describe the articles encountered. The number of heading were limited until recently. Between 1966 and 1980, we did not have the heading "clinical trials". We used instead "clinical research". However, we did have a number of terms describing study characteristics, retrospective studies, follow-up studies, and statistical methods.

There was little change between 1980 and 1989. Since 1990, we have added over 40 terms related to clinical research and epidemiology. I will highlight only a few. The first group relates to study design and include basic descriptors, such as double or single blind studies and random allocation, common to most trials. The second group includes the various types of studies, such as case-control studies and prospective studies, that are used in conducting trials in epidemiology. The third group comprises factors that may be under study or affect the outcome of a study. These include

the various biases such as placebo effects. We have also created a grouping for evaluation studies that includes a set of terms for the various classes of trials themselves and the drug approval process. Finally, we have added a few headings related to outcome and process assessment which are important in health services research.

In 1991, we also created a new concept called publication types. These headings are used not to describe the subject content of articles, but to describe the format in which the material is presented. There are about 40 publication types, including clinical trial, randomized controlled trial, and multi-center study. Every article that we identify as the report of a clinical trial receives one or more of these publication types, in addition to headings describing the characteristics previously described and headings pertinent to the subject under study.

The institution of publication types is a perfect illustration of database fluidity and our practice of maintenance. We have maintained every citation that we could identify back to 1966 by adding one or more of the clinical-trial-related publication types. This enables retrospective retrieval to be performed on the same basis as prospective retrieval and should improve the recall statistics. As I mentioned with the journals, please contact me with new headings that would be important in the characterization of clinical trial data. I would like the vocabulary to be perfect. The key to that, however, is how well it serves the user. A vocabulary is no good unless it is clear and straightforward in its presentation and accessible to the user, no matter how organized. Feedback is very important.

In closing, I would like to review the steps we have taken to improve our clinical trial coverage and to briefly describe our commitments for the future. I have briefly described the indexing policy modifications and the vocabulary expansion that has taken place since 1990. In addition, we have established a monthly review mechanism, to scan citations containing expressions of clinical trial characteristics in titles or abstracts, review those that have not been indexed with a clinical trial heading, and to add the appropriate headings when necessary through database maintenance. This process can also serve to identify potential new terms for addition to MeSH.

We have developed a search strategy to optimize retrieval of pre-1990 clinical trial literature. This strategy will be evaluated and tested and, when approved, we shall make it available, probably as a fact sheet to any interested parties. We continually speak with editors and publishers about the need for explicit characterization of clinical trials to maximize retrieval potential. And we shall continue to pursue these avenues of approach. Finally, when a registry may be established, we would commit to the linking of published trials in the registry with companion MEDLINE citations and to ensure that every linked MEDLINE citation would be characterized by the relevant clinical trial heading and publication type.

I hope that this presentation has given you not only a picture of indexing practice and database coverage of clinical trials at the NLM, but also an understanding of the commitment we have to ensuring the most accurate and complete representation of the literature possible.

SEARCH STRATEGIES FOR FINDING RCTs IN MEDLINE

Carol Lefebvre
UK Cochrane Centre
NHS R&D Programme
Oxford

DR. LEFEBVRE: The Cochrane Collaboration is an international organization chaired by Professor David Sackett interested in preparing, maintaining, and disseminating systematic reviews of the effects of health care. Systematic reviews, in order to be reliable, must consider all the relevant evidence. From our point of view, sensitivity is more important than specificity. More specific criteria can be applied subsequently by investigators preparing systematic reviews. These points have also been mentioned today by Kay Dickersin.

Dr. Dickersin, Dr. Scherer, and I have prepared over the last few months a systematic review of twelve studies on MEDLINE search strategies for identifying RCTs which will be published in BMJ. I will present some results of that study, some of which are related to the results of the ophthamology studies presented earlier by Dr. Dickersin.

The percentage of RCTs identified by MEDLINE searches was determined. RCTs were identified for a series of investigators studying a good range of biomedical topics. The percentage of RCTs identified in MEDLINE searches was presented. The baseline was all RCTs identified by searching MEDLINE and other databases, hand searching journals, writing to colleagues, and other methods of information retrieval. There is quite a range: 17% is the lowest, 81% is the highest, but the mean is 50%. Therefore, on average, in these twelve studies, only 50% of all available RCTs can be identified through a search of MEDLINE.

This information is of interest to the clinical community and reviewers, in particular, because they want to know the amount of available RCTs they have identified.

We have also determined the percentage of RCTs identified by a MEDLINE search of all relevant trials in MEDLINE. This was a re-evaluation of the figures used to determine the previous percentage. It identifies the number of RCTs actually in the MEDLINE database that can be found by a MEDLINE search. This should be 100%. But, the range for this percentage goes up to 91%. In some cases, it is a low figure. The average is 78%.

Some studies approached this problem in a slightly different way—instead of using the complete MEDLINE database, they used selected journals. The average percentage of RCTs identified by MEDLINE searches of selected journals for several investigators was 64%. The figure should be

higher, because in some cases the subject element would not be lost, particularly if only the journal title is searched.

There are several reasons for poor sensitivity of MEDLINE searches for RCTs. First, if the authors do not explicitly describe the research design, the National Library of Medicine (NLM) indexers cannot determine sensible indexing terms. Editors also need to have guidelines to ensure that the methods are explicitly stated. Second, suitable indexing terms have not always been available or applied consistently. Finally, the appropriate search strategies need to be used. MEDLINE is a very structured database. If the appropriate search structures are not used, then the results will be poor.

These concepts were presented by Peri Schuyler. However, I would like to credit NLM for their work. NLM has responded well to the user demand for better identification of RCTs. Between 1966 and 1977, a specific term was not available for RCTs. In 1978, the term random allocation was added as a descriptor, and those items where the method was clear could be indexed under this term. In 1990, randomized controlled trials was added as a medical subject heading, followed by randomized controlled trial as a publication type in 1992. There have been great steps made by NLM to cure this problem.

NLM is further advanced than Elsevier in this regard. Elsevier, as Kay Dickersin mentioned, produced the EmBASE database which is the main competitor to MEDLINE. It is an international database that concentrates on European material. It is widely used in Europe. Until this year, they had no suitable way to identify a randomized controlled trial; they could only identify clinical trials. As a result of a meeting we had with them in January of this year, they will be introducing a whole range of methodological indexing terms. They are indexing to those terms now, and the updates to the database from January next year will contain the new terms.

In developing a MEDLINE search strategy for identifying RCTs, from our point of view, sensitivity is more important than precision. We held an international workshop on building a register of RCTs on the eighth of November, 1992. We invited experts who were experienced in searching for RCTs on MEDLINE. Many had published their work, such as in the twelve studies I spoke of earlier. Others had not published but had made good progress in identifying and solving the problems. We compared alternative search strategies and selected a strategy with optimal sensitivity for identify as many of the RCTs in MEDLINE as possible. This will be published in *BMJ*.

This search strategy has 34 lines (see Fig.1) and is divided into three sections according to the likely precision. The convention in this figure is capital letters hyphenated for the indexing terms and lower case for the free text terms. This varies with the software. The first phase contains terms like randomized controlled trial as a publication type. It should be precise. The final, penultimate search line removes the material which is solely animal, but retains the human material, or the material not

Designing a search strategy for optimal sensitivity

#1 RANDOMIZED-CONTROLLED-TRIAL in PT
#2 RANDOMIZED-CONTROLLED-TRIALS
#3 RANDOM-ALLOCATION
#4 DOUBLE-BLIND-METHOD
#5 SINGLE-BLIND-METHOD
#6 #1 or #2 or #3 or #4 or #5
#7 TG=ANIMAL not (TG=HUMAN and TG=ANIMAL)
#8 #6 not #7

#9 CLINICAL-TRIAL in PT
#10 explode CLINICAL-TRIALS
#11 (clin* near trial*) in TI
#12 (clin* near trial*) in AB
#13 (singl* or doubl* or trebl* or tripl*) near (blind* or mask*)
#14 (#13 in TI) or (#13 in AB)
#15 PLACEBOS
#16 placebo* in TI
#17 placebo* in AB
#18 random* in TI
#19 random* in AB
#20 RESEARCH-DESIGN
#21 #9 or #10 or #11 or #12 or #14 or #15 or #16 or #17 or #18 or #19 or #20
#22 TG=ANIMAL not (TG=HUMAN and TG=ANIMAL)
#23 #21 not #22
#24 #23 not #8

#25 TG=COMPARATIVE-STUDY
#26 explode EVALUATION-STUDIES
#27 FOLLOW-UP-STUDIES
#28 PROSPECTIVE-STUDIES
#29 control*or prospectiv* or volunteer*
#30 (#29 in TI) or (#29 in AB)
#31 #25 or #26 or #27 or #28 or #30
#32 TG=ANIMAL not (TG=HUMAN and TG=ANIMAL)
#33 #31 not #32
#34 #33 not (#8 or #24)

specified to be human. Some elements are not indexed as human in MEDLINE but are of interest to clinical trials work.

The second phase contains terms with moderate precision, including free text terms. Terms such as CLINICAL-TRIAL in publication type, clinical trials as a medical subject heading, PLACEBOS, RESEARCH-DESIGN. Followed by a range of free text terms expected to identify randomized controlled trials.

The third phase utilizes terms with low precision. For example: volunteer*, truncated, or control*, truncated. The asterisk is the truncation symbol in this particular software. These terms yield a lot of irrelevancies, however, we want to create a registry with as many necessary terms as possible.

I have developed a technique to assess the precision of individual terms with a colleague at Oxford, Clive Adams, a psychiatrist with expertise in the field of mental health RCTs in MEDLINE. We developed a method to check the precision of each individual term by running the search once for each set number, and on each occasion, omitting a single term. This identifies the items added by including that term. This will be published in *Psychological Medicine*.

This strategy will be used by Kay Dickersin and her colleagues at the University of Maryland to create a core registry by scanning abstracts and attempting to identify RCTs. This project has been funded through a contract with OMAR. It was planned many months ago, and we hoped we would have a small database to show you. Unfortunately, there have been problems with equipment delivery. Thus, I proposed an alternative—an up-to-date assessment of MEDLINE in 1993. NLM has implemented many improvements.

I designed a search strategy to identify reports with random*, truncated: random, randomly, randomized, and so forth—random as a text word in the title or the abstract on the MEDLINE database—but excluding those articles which have been indexed by NLM with RANDOMIZED-CONTROLLED-TRIAL as a publication type, or randomized controlled trials as a medical subject heading, or RANDOM-ALLOCATION. I wanted to determine whether, in early 1993, there are still RCTs in the MEDLINE database that we could not identify using the indexing terms. These should be indexed with the term RANDOMIZED-CONTROLLED-TRIAL as a publication type. Peri Schuyler explained the purpose of that term earlier. I also allowed RANDOMIZED-CONTROLLED-TRIALS and RANDOM-ALLOCATION as medical subject headings options as well, because I think it is reasonable to include those terms.

This search was run for the first six months of data for 1993 (NLM update codes 9301 to 9306). These reports are not necessarily published in 1993 but are indexed during those months for the hard copy of *Index Medicus*. In January, there were 409 references that were retrieved by this search

strategy. Students scanned a total of 2,411 titles and abstracts from the MEDLINE database identified by this search to identify those which were definitely stated to be RCTs.

There are two reasons why I say "definitely stated to be." First, we are interested in sensitivity and want to include items which may not be RCTs. For the purpose of this study, the definition was quite different. To be classified as an RCT, it had to be stated in the title or abstract that it was definitely an RCT. Second, even if the author states it is definitely an RCT, when the research is further examined, we may disagree. One can not comment on this within the MEDLINE record.

There were 432 items which the students considered to be definitely stated to be RCTs. Over the six months evaluated, the percentage varies. The average is 17.9%. The percentage is less relevant in many ways than the number, because this is a quick way to find unindexed RCTs. If other terms were used in the search strategy, different items would have been identified.

To determine if there were changes since this study was completed, we searched the November data using the same criteria. We found 52 references to RCTs; this was 17% of the total references found. This was the most current data available on MEDLINE.

To determine how well the students searched, I read the 96 references from January. For 5 of the items, I need to consult with a more knowledgeable expert. For the remaining 91, I definitely agree with the students' interpretation.

I will discuss samples listed in a database of approximately 430 records from PROCITE. The first item is from the MEDLINE record of an article in the *New England Journal of Medicine*. In the abstract it says, "in a randomized, multi-center trial..,". The second quote from the *Lancet* reads, "of these, 26 agreed to be randomly assigned to medical treatment ... or revascularization." These samples illustrate that NLM indexes core material correctly because we assign our best indexers to that and concentrate on those areas instead of more obscure journals. These are examples of what are generally regarded as core, very important international journals.

The next one was chosen because it has a good clue in the title—it says double-blind. We cannot rely on this alone. However, in the abstract it says, "Patients were then randomly assigned to receive acetyl levocarnitine ... or placebo". I could have chosen any of the 430 studies; none stood out. These were chosen because I thought they would be of interest to the members of the audience from the National Institutes of Health. The first, from *Fertility and Sterility*, relates to women's health, "embryos . . . were randomly allocated to a control or biopsied group . . ." The next, from *AIDS*, said in bold, capital letters, "DESIGN: Double-blind, placebo-controlled, randomized clinical trial..." This one, as many others, had a structured abstract. For those of us who have thought that structured abstracts would improve indexing, this is discouraging. The example from the *Journal of*

the National Cancer Institute reads ". . . eligible patients were ... randomly assigned to receive or not to receive four courses of. . ." .

Belfort, et al. in the *American Journal of Obstetrics and Gynecology* said, "The effect of magnesium sulfate on maternal retinal blood flow in preeclampsia: a randomized, placebo-controlled study". This is atypical. Of the 2,411 references that the students reviewed, only one that was not indexed correctly had "randomized controlled trial" in the title. Probably 100 or 200 clearly stated "randomized" in the abstract. An indexer at the NLM could be proud of one wrong in 2,411 articles indexed. However, what happens between the titles and the abstracts? This is something I have recently realized by doing this study.

To put MEDLINE in context, if we consider all RCTs as our goal, they can be divided into completed and ongoing. MEDLINE indexes reports of completed trials. Other colleagues at this conference will discuss ongoing and planned trials. Within the completed trials, there are published and unpublished trials. Other colleagues will discuss completed trials which do not become published.

Of those that are published, some are published in journals. Others are published in other literature, commonly referred to as "gray literature", abstracts and theses. (We are not discussing this over the next couple of days, but I think it is an important issue.) Within journals, some are indexed by MEDLINE, and some are not. MEDLINE indexes 3,700 journals, and in the NLM collection alone there are 16,000 biomedical journals. Thus, 12,000 are not indexed. In addition, MEDLINE began in 1966, and RCTs were performed before 1966. Within MEDLINE, certain RCTs can be identified, but a large number cannot be.

In conclusion, MEDLINE is the best bibliographic database available for identifying RCTs. In addition, it is currently the best source available for identifying RCTs. In spite of this, however, it remains difficult to identify the high proportion of clinical trials which are necessary for informing evidence-based health care

DISCUSSION

DR. FERGUSON: It seems to me that in addition to problems in definition—it seems like the definitions have, even from the same individual, been different at different times, for example, Dr. Meinert's definitions—that the definitions used by those looking through the literature have been different. The National Library has amended their definition as time goes on.

I would like to open up the floor to questions and comments.

DR. HUTH: Could Peri Schuyler tell us whether the terms she refers to are all on one piece of the tree, or are they scattered in the MeSH tree?

DR. SCHUYLER: The bulk of them are in a tree called N5.

DR. SIMES: I think one purpose of this workshop is to develop some recommendations for ourselves, governments, or journals. I think one of the useful recommendations for assisting MEDLINE is to add an additional step to the review process of clinical trials by journals, of classifying them as randomized or not. For example, most journals now have a statistical review, perhaps that could include not only that the article is worth publishing but also that the trial is classified as a randomized controlled trial. If that becomes part of the review process, and those terms are used systemically, that might help MEDLINE later.

DR. HORTON: I am from *The Lancet*. Perhaps I could address that point. Regarding published randomized controlled trials, I would like to emphasize Carol Lefebvre's point about the role of editors in aiding search strategies. Using Kay Dickersin's words, editors also need to be more careful on this particular point. In many ways editors are in the best position, because they represent the final pathway of publication.

Perhaps one recommendation should be to seek the cooperation of international medical journal editors, perhaps through a representation to the Vancouver Group which meets in Oslo next year. If the recommendations from this workshop were submitted to them, I am sure they would go on their agenda.

DR. FERGUSON: In effect you are saying perhaps that the editors should be responsible for these determinations?

DR. HORTON: Ideally, yes. They represent the final path, and authors are not always aware of these issues; whereas editors, who should be at this sort of workshop, are.

DR. HUTH: (an editor) I agree. It seems to me that appropriate and relevant MeSH terms might be a basis for a list of terms from which authors could be expected to characterze their report from their point of view. This could be scanned by the editor and would help force authors to use more explicit descriptions. I agree that editors have a responsibility for this. I think they could

improve reporting better than authors could. Authors often do not read information from author's information pages.

DR. ORZA: I am from GAO. One thing to consider while deciding how inclusive or exclusive we want our definition to be, is the purpose or purposes of the registry. If we want to collect all research in one place, then we want a more inclusive definition. But if we want the registry to encourage the kinds of research we would like to see, and discourage the kinds of research that we do not like to see, we would choose a more exclusive definition.

DR. MOHER: I think that we ought to rely more on authors. Almost all journals I am familiar with have instructions to authors. Part of that instruction could be to state in the title if the authors' have performed a randomized trial. Although editors have an important role, I think authors have the most important role for doing that. Perhaps this could be discussed by the Vancouver Group.

DR. BALAS: I am from the University of Missouri in Columbia. One of our projects is collecting randomized controlled clinical trials and testing computer assisted information intervention. This is a special area of health services research, and I would like to highlight the problem of the unit of randomization. The unit can be the patient, the encounter, the physician, or groups of physicians. Sometimes authors, for example, report randomized controlled clinical trials when one hospital randomizes the intervention group, and another has randomized its control group. I do not think that is a randomized controlled clinical trial. In the case of developing a definition, I recommend defining the unit of randomization as well.

DR. LEBRON: I am from the *Online Journal of Current Clinical Trials*. I have a two-part comment. From Dr. Dickersin's presentation, six of the materials that she studied, quite a large proportion of them, had information in the title and in the abstract pertaining to being a clinical trial. With this as the basis of information and considering that indexers are human beings, when these data are initially entered into the database (after all you are dealing with hard copy journals and someone has to keyboard this information for MEDLINE usage) is it possible to have a small set program that could do a quick indexing automatically, and then have the indexer indicate whether or not the proportions of relevance of the different items that have come up are correct? I think in that particular case we would increase the possibility of having the indexers identify more randomized clinical trials. We should remember that indexers get fatiqued from looking at these repeatedly. Therefore, some computerized assistance might minimize the problem.

DR. SCHUYLER: I think it is a very interesting idea, and one that we might look forward to experimenting with on a small scale. I think that in part, while indexing, the abstract is not the first line of approach. The first line of approach is the author's title and the stated purpose of the study. It is only after the indexer has reviewed the contents that he returns to the abstract. So while not

wishing to make an excuse for the indexer, I think that it should be recognized that the abstract is not the first thing the indexer examines.

DR. MARIN: I am from the Emblematics Corporation. We have focused on the indexing, tagging, and labeling. What about searching techniques? Suppose the term randomized was not there, but many terms that pointed to a randomization were? If we could not identify the material in the first place, but had an alternative logic, such as signal detection theory, we might be able to identify things that were not stated implicitly in the context. I would suggest this type of searching might be added to the currently avaiable kinds of techniques.

DR. LINDBERG: I think that is a good suggestion, as were the earlier ones. Actually we have been experimenting for years with programs to automate indexing. We are aware of the ideas, but they do not work well enough to replace human indexing—largely for the reasons you mentioned, that the indexers are able to pick up ideas that are expressed, or index by words that do not occur in the publication. While we know about the power of free text searching, and use it on the Internet, it is practically the only method used. We think there is probably a good role for human indexing in the foreseeable future. As for your ideas about the registries, all of those things are necessary. I do not think that any one of these will substitute for the other.

I guess we should mention Unified Medical Language System, because that is yet another level of dealing with the semantics. The index terms in MeSH are part of a much larger group called Unified Medical Language System, which includes approximately 300,000 terms including the fixed vocabularies from the International Classification of Diseases and a number of others. The purpose of that project is to understand medical meaning and to look toward knowledge bases that will allow programs to reason their way to accomplish the things that you are talking about. I do not think any one should be an excuse for inactivity on the others. Your registries would be wonderful if you can get them done.

DR. MARIN: I would like to provide feedback on that comment. Basically, there is not one strategy but a combination of strategies. This is not saying that that a human should be replaced, or a machine is better, but an interacting combination might lead to a kind of synergy effect.

DR. SCHULZ: I am from the Centers for Disease Control. While we are making pleas for authors to say that they have done randomized trials, I would like to make one further plea, that the authors state what they have done. As known from our research, a lot of authors who state that they have done a randomized trial have not done one. In many trials that may have started out as randomized trials, the actual process is not concealed (completed). So, what we think may be an unbiased study really is not. Instead of encouraging people to say randomized or not, we ought to go

one extra step and make them include information on how the allocation schedule was developed and how it was concealed.

DR. FERGUSON: Would it be possible, Dr. Schuyler, if words were used in a title or in an abstract that said random allocation when, it was something else, that an indexer would pick that up?

DR. SCHUYLER: I think we need to emphasize that the indexer will not interpret. The only thing that the indexer can do is to read the content, understand the content, and not to determine whether the author was making [illegible] not. I think it is beyond the scope of what should be expected from an indexer.

DR. HAIGLER: I am from NIMH. It sounds like the role of making sure that it is a randomized clinical trial is with the editors and the reviewers. When reviewing a paper, the reviewers should be instructed by the editor to make sure that, if someone says it is a randomized clinical trial, that their methods reflect it. If not, it should be rejected. Or, if too little detail is given, more detail should be requested in the paper.

DR. SIMES: That was the point I made earlier. I think it would be useful if the reviewers classifying studies as well as approving them.

I was going to suggest an additional term for consideration for indexing trials in the future, "prospectively registered". We are discussing establishing registries of published trials. We would like to have as comprehensive a set of studies as possible which we are searching for from the published literature or from studies submitted for publication. Later on, if we are doing meta-analyses of trials where there may be unpublished studies, it would be useful to know which of those studies are contained in a registry. Thus, I advocate adding whether a study is prospectively registered with a registry of trials as an extra field.

DR. LINDBERG: To follow up on that, we could consider a very heavy-handed approach—and I am not suggesting this, I am only describing what is at the end of the rainbow. In the case of molecular biology, top journals will not accept a paper unless the sequences have already been deposited in an appropriate computer database—GenBank for nucleotide sequences and PIR for amino acid sequences. Thus, in these journals, the article will not accepted unless a number from the registry is included.

This heavy-handed approach works surprisingly well, in part because all business is done electronically. ACTGs are not actually typed in very much anymore. They are transfered from one computer to another by electronic mail through Internet. The submission, issuing of a number, and confirmation all occur rapidly, and do not seem to present as much of a problem as we expected. I suppose if it was agreed internationally that this is what should be done, it is technically possible to implement that kind of very tight coupling.

DR. FERGUSON: It is a good point. I think that we will probably talk more about that as we get into the registries of ongoing trials, and the linking that the Library could and probably would do, as Peri Schuyler mentioned.

DR. WAGNER: I am from the University of Pittsburg School of Medicine. In your opening remarks, while stating the meeting goals, you said some things which do not fit with some of the presentations. I think you said we were going to exclude completed trials in the registry.

DR. FERGUSON: No, I was saying that a registry of ongoing trials would be a major focus of later this morning, this afternoon and tomorrow. I said that the registry of ongoing trials would not include results of the trials.

DR. WAGNER: The general consensus is that we are going to be talking about the design of a registry of ongoing trials?

DR. FERGUSON: Yes.

DR. WAGNER: Or about a registry of completed trials?

DR. FERGUSON: Yes, I think both. Completed, but not yet published. I was separating trials for matters of discussion.

DR. WAGNER: From the information-needs perspective of people doing cost effectiveness analyses and clinical decision analysts, a set of completed clinical trials is more valuable. Thus, I was wondering about the value of focusing on uncompleted trials.

DR. FERGUSON: I guess I do not want to make that a false statement. I was distinguishing between once a trial is completed but is still not published, it would presumably be in the registry as a completed trial. We would hope to have information such that the randomiztion had stopped in the registry of ongoing trials, even though it is completed. Once the trial was published, one would hope that that is now linked to a published document. I do not mean to exclude completed trials in the ongoing registry.

DR. WAGNER: You also said that we were not going to include the results. Again, if we are going to talk about systems and approaches, we have to consider the results of the clinical trial. Also, from the perspective of an information consumer, I must disagree; I think that it is very important. For example, I am a clinical decision analyst, among other things, and most of my comments at this meeting are going to be from that perspective. When I use the literature, ideally the probabilities that I need are presented. So if I use a registry of clinical trials, I would like the probabilities, or some description of the probabilities that I expect to find as a result of the trial, to be given. Which leads to the issue of what should be represented in this registry? Should it be content-free? Or should it contain a lot of content and explicit representation beyond the simple indexing level? These are some of my needs as a consumer of such a clinical registry.

DR. GRAY: I am Director of Health Policy and Public Health and also the Oxford Regional Health Authority. I am currently working with Iain Chalmers to develop a registry for surgery and anaesthesia trials. The prospect of adding details to the registry is mind-boggling. It is a big job to find the trials. The previous discussion about the number of non-MEDLINE journals illustrates the size of the exercise. I think our priority has to be to obtain the citations and, perhaps, the abstracts. Collecting the data should be left to people who want to use the registers.

DR. LEFEBVRE: I think there has been a certain degree of confusion because of the opening remarks about ongoing trials, which have led some of the audience to think that our meeting is not going to address the need for a registry of published and completed trials. I think that might be the root of the confusion. I think the gentleman at the back may have thought that was what you meant when you said we were going to be concentrating on ongoing. I am not sure whether that was what you wished to say.

DR. FERGUSON: I think that the talks that we are having today and tomorrow seem to be more focused on the registry of ongoing trials. I am not excluding the other.

DR. LEFEBVRE: We are not assuming that there is no need for a registry of..

DR. FERGUSON: I assume nothing.

DR. TEMPLE: With the last group of speakers, I am growing increasingly confused. I am from the FDA. As I understand it, there are two quite discreet questions being asked. One is how to assemble data that is already in the public domain and in print; which is much harder than it sounds, because abstracting is difficult and because techniques of finding words like "double-blind" in an abstract do not seem to be available. Howevr, that seems a manageable problem. It is difficult, but one knows how to do it. We begin with everything in print (or found in MEDLINE).

The matter raised at the beginning was whether we could try a more difficult thing—to identify trials that are not in print, but that someone has begun to undertake. That is not very hard when they are NIH trials, because presumably you know about them. But when they are funded through a variety of sources and governments, that is a formidable challenge. Data will not be available because the results of the trial are not yet known. It is a way for people conducting trials to what others are doing so they can talk to them, not bother, or do other things. I guess you have to separate the two things because they are totally different endeavors. It isn't either/or. They each have different uses.

DR. FERGUSON: I agree with you. That was one of the things I was trying to say in the beginning.

DR. T. CHALMERS: In evaluating the efficacy of MEDLINE searching, I am surprised the discussion has not turned more towards the advantages of structured reports of clinical trials. We have

had a good example of its usefulness regarding abstracts. I also do not think there has been enough emphasis on editors and investigators developing plans together for structuring clinical trials so that the indexing will be more efficient and easier.

A second thought, regarding the registry of unpublished trials. I think the best way to kill this project is to discuss including results. That is a quantum leap towards an impossible accomplishment.

DR. EVERHART: I am from the NIH. There are some terms slashed by Phase I, Phase II, Phase III and, Phase IV, that NLM uses. In Dr. Pinn's remarks, she mentioned that NIH will be challenged to come forward with all Phase III trials. I wonder if these terms should be defined in the discussion here, particularly Phase IV.

DR. SCHUYLER: They are post-marketing studies.

DR. TEMPLE: Those are our terms, so I will comment. The terms are useless for this discussion. They are descriptions of convenience used in our regulations to describe an orderly process of drug investigation.

The question here is whether the trial is randomized, controlled, and has the characteristics of a trial you want to find. It does not really matter whether the drug is already marketed in this country or another country. It does not matter whether it is called a Phase II trial because it is the earliest controlled trial of a drug, or a Phase III trial because it is a later. The characteristics of the trial are what matters.

The one exception that I think Larry was trying to make was that the kinds of trials where endpoints like mortality are examined are generally called Phase III trials, although technically they are often Phase II trials for some drugs. It is not useful terminology for this purpose.

DR. FRIEDMAN: Except that it has led and continues to lead to confusion, because people still call them clinical trials when by my definition, and I think the definition that most people here are using, they are not clinical trials. I think that is where it is harmful.

DR. TEMPLE: What isn't a clinical trial?

DR. FRIEDMAN: A Phase I clinical trial, for example, does not meet my criteria, typically, of being controlled.

DR. TEMPLE: Some of them are not what you would want to put in a registry.

DR. FRIEDMAN: Right, but if the title or the abstract says "clinical trial", and it is listed that way, then it causes confusion.

DR. TEMPLE: But we even consider as clinical trials an open safety study with no randomization. You can use any definition you would like.

DR. FRIEDMAN: Not my definition.

DR. CHARPENTIER: I would like to stress that randomized controlled clinical trial is probably the best facility around for therapy trial. However, the importance of randomization should not be overstressed for, for example, diagnostic procedures; randomization with triple arms is probably stupid. Self-controlled trial such as a crossover or partial crossover is probably much better. If a registry is to be established on such procedures, special key words describing the methodology should be made according to the subjects. For example, for one example of a strategy of researching MEDLINE that was quoted, I was stricken by the fact that the word random was probably overstressed.

DR. ORZA: This relates to my earlier comment. I would prefer that people not do unrandomized or historically controlled or uncontrolled trials. But, if they are going to do them, I would like to know about them. It sounds to me that we have already decided about the kinds of things we are going to include in this registry, and some of those things that might be left out. Am I misunderstanding?

DR. FERGUSON: You are misunderstanding. That has not been decided.

DR. ORZA: So we are still trying to decide on the definition of what goes in?

DR. FERGUSON: I think we are.

DR. ORZA: I think one of the best arguments for including those other things is that you can then study them and compare—for example, do the ones that we think are better, get different results?

DR. FERGUSON: That is why we are here.

(At this point a brief recess was taken, after which the meeting proceeded as follows.)

THE QUALITY OF LEVELS OF EVIDENCE FOR HEALTH CARE DECISIONS I

C. Hudgings

Agency for Health Care Policy and Research

DR. HUDGINGS: I want to cover four major topics. The first is to explore why the Agency for Health Care Policy and Research and, specifically, the Office of the Forum for Quality and Effectiveness in Health Care (the forum), which develops clinical practice guidelines, would have an interest in the topic that is being explored at this conference. I want to discuss some key points about guideline development and presentation of guideline statements. I want to then describe two challenges that we face in examining the scientific data and other information and formulating guideline recommendations that dovetail nicely with the challenges you have in front of you today. Finally, I will pose some questions that may act as a framework for the future in exploring these two issues. That particular information comes from a conference held November 1 and 2, 1993, at which the Agency convened experts to look at the various levels of evidence and how they support clinical practice guideline statements.

Dr. Clinton gave a good overview of the purpose of the Agency for Health Care Policy and Research (AHCPR), but I want to give again the definition of practice guidelines that was prepared for us by the Institute of Medicine. I call your attention to several different aspects of this definition. One is that the guidelines produced under the auspices of the AHCPR are systematically developed statements designed to assist two primary groups, practitioners and patients, in making decisions about specific clinical circumstances and conditions.

The methodology that AHCPR has advocated in the development of its clinical practice guidelines is very specific and has many different distinguishing characteristics. I call your attention to these five. The methodology for guideline development is explicit and systematic. It is all laid out. We give it to the panel chairs, the methodologists, and the other consultants who work with the panels in developing the clinical practice guidelines. This systematic development methodology calls for the use of a particular analytic framework, that is, the method by which the literature and other information are analyzed to derive the practice guideline statements. It involves a systematic look at all of the available science that is relevant to the particular condition under study, and an analysis of the benefits and harms of the various interventions for that particular condition. It is evidence-based in that science is used whenever possible to drive the formulation of clinical practice guideline statements. However, expert opinion is also used in deriving the statements. Expert opinion also comes into play in the interpretation of the literature, as you will see in a minute. Finally, the

methodology and processes for guideline development are subjected to documentation, so that the reader can tell exactly what the panel's thought processes were.

It is important to realize that a clinical trials registry, whether it contains ongoing information or information about completed trials, would be of use to AHCPR at three different points in time. One is at topic selection—when we are considering various topics for guideline development. The second is when we are developing a guideline. Third is when we are updating a guideline, as we are now for the six guidelines that have been published by the Agency.

Following is a typical presentation of the guideline statement. The latest guidelines that we are producing, and those that are in process, generally have this kind of format. There is a sentence or two sentences labeled guideline or recommenda[illegible]hat give the general essence of the guideline statement. Generally, this is followed by a paren[illegible]cal comment that reads strength of evidence equals "something". That "something" is one of the particular challenges that I want to discuss in a little bit.

Beneath this guideline statement is supporting information, which indicates such things as the justification and the rationale for the guideline statement, the key studies that were used in deriving the statement, and whether the panel is generalizing from the study populations.

To clarify that typical presentation, I have some examples of practice guideline statements taken from the depression guideline. This particular statement speaks to whether or not a second medication should be used in addition to the initial medication. You will see that the strength of evidence is labeled "A" for this particular guideline statement. This is another guideline statement derived from the depression practice guideline, for which the strength of evidence is labeled "B". There is a range in the types of evidence that support the different guideline statements.

We know that there are many challenges that confront us in identifying the relevant literature for the condition under study, obtaining copies of the articles, reviewing and analyzing the studies, and pooling the evidence to help us come up with a recommendation in a particular area. The scientific literature, and other information that we use to assist the panel in shaping a guideline statement, generally comes in several forms. We are not restricted to the use of clinical trials or randomized clinical trials. That would make our clinical practice guidelines very short, because we have found that there is a dearth of literature to support several of the recommendations that we would like to make for the conditions under study.

I spoke earlier of two challenges that are particularly relevant to this topic that were the subject of the recent conference. These challenges are: How do we rate and describe the quality of the literature that is surveyed to address the particular research questions? And, how do we describe and classify the strength of evidence for a guideline recommendation? Both of these components are

important for users of the guideline to help interpret the guideline statements to see how they fit within their particular practice.

Our experience with the guideline panels today has shown that there is some variation in how the panels have approached these two activities or tasks. I would like to give you some examples of that panel experience.

In terms of the quality of the literature, the panels generally define a number of key variables about the study design that are used in rating the quality of an individual study. These variables include such things as patient selection, sample size, aspects of the method, and design of the study. They are many of the characteristics of the study that we have been describing here today.

Some panels have chosen to rate studies with very broad parameters, such as poor, fair, or good. Some use a more systematic method for rating the quality of an individual study. For example, one panel identified eight particular criteria, rated each of those eight criteria on a three-point scale, and then only included studies that scored 15 or higher out of the possible 24 points. Generally in looking at the quality of the literature, the panels use a hierarchy of study design much as has been presented earlier this morning: meta-analyses, randomized controlled trials in a multicenter situation, randomized controlled trials in a single institution, cohort studies, case control studies, and observational studies.

Another method that a panel has used is very explicit with four different categories and subsets of some of the categories. For example, the Acute Pain Guideline Panel classified the evidence according to meta-analysis, randomized controlled trials, controlled studies without randomization, quasi-experimental studies, non-experimental studies, and expert opinion.

The second major challenge that we discussed at the recent conference was how to describe and classify, or do we need to classify, the strength of a guideline recommendation. That particular label is very helpful to users in knowing how much weight to give to a particular recommendation. Within broad parameters, our panels have used several different methods to describe the strength of evidence. Many of these methods use the U.S. Preventative Services Task Force designations as a starting point. One panel used three designations and said that A-level strength of evidence is good research-based evidence, B is fair, and C is no research-based evidence, but evidence based upon expert opinion and panel consensus. As you may see, problems arise in defining the good and fair aspects of these definitions. Another panel chose to use labels in describing their strength of evidence that might fit more with the David Eddy model of such description. For interventions, they have said that they would label the intervention as a standard. That is, evidence clearly shows that the intervention is superior to its alternatives, that the intervention should be labelled as a guideline, that the evidence points to the fact that the treatment is probably superior to other kinds of alternative interventions, but

the final vote is not in. Or, perhaps, that the intervention is an option, and we really don't know the outcomes with certainty.

Another panel described seven different levels, and the levels related to the quality of each individual study. Then they collapsed those levels to get three designations for the strength of evidence, so that A-level evidence was evidence from well-conducted RCTs or cohort studies, B was fair evidence based on other types of studies, including observational studies, and C related to expert opinion.

The final example is a panel that had access to much randomized controlled trial literature for their condition. As you can see, they restricted the A-level designation to randomized controlled trial literature, B-level evidence to other kinds of literature, and C-level evidence to no clinical studies, but rather evidence based on panel consensus.

What was the discussion at this conference, and what directions are clear for us in exploring these two topical areas? Our goals for defining parameters for the panels to use are that whatever mechanisms we come up with must be simple, easy for the panel to use, and easily understood by the users of the guideline. Everyone who reads the statement and any associated information should come to the same conclusion and should understand what is meant by the labels or words that are applied. There must be organized flexibility in the approaches used, meaning that the methods must be adaptable for different kinds of literature, because it is clear that the quality rating scale that one might derive for that of randomized controlled trial literature may not be the same as that derived for other types of literature.

The conclusions of the conference were that the strength of the evidence includes several broad parameters which must be considered in deriving a classification scheme for that level of evidence or strength of evidence. One factor is the strength of the evidence, including not only the quality of the studies, but also the consistency of the study results over time, across populations.

A second major characteristic to consider is the relevance of the evidence, and this applies in large part to the generalizability of study findings. We found that RCT data are often quite limited in terms of the populations that are covered vis-a-vis the populations of interest for the total guideline development; a decision must be made about how far the results of any particular clinical trial can be generalized to the population of interest in the guideline.

Two additional factors to be considered are knowledge about basic science and basic principles of practice, because there are some interventions that are so obvious that they have never been and never will be subjected to clinical trials. Yet the use of those interventions in providing care to the typical patient by the average practitioner would be generally accepted.

Our goal is to get guideline statements that are meaningful to users. We must take a multidimensional look at the quality of the literature, the strength of the evidence, and the strength of the recommendation. It is very likely that whatever scheme we evolve over time will be different for the different types of literature that we find support the different guideline conditions under study.

THE QUALITY OF LEVELS OF EVIDENCE FOR HEALTH CARE DECISIONS II

Gordon Guyatt
McMaster University

DR. GUYATT: When developing treatment recommendations, you must look at a body of evidence. The first step is to carefully specify the question according to the population, intervention, and outcome. Two quite separate issues arise when assessing the strength of evidence: the methodologic quality of the evidence and the magnitude of effect. These issues have, to some extent, been confused. Finally, one must draw an inference, make recommendations, and summarize the levels of evidence.

The process of determining levels of evidence has been evolving. We now know that methodologic quality and results must be separated. If you think of sample size not as an issue of methodologic quality but rather as an issue in the precision of the results, it is clear that there is a degree to which any given treatment works. Further, statistical significance and clinical importance must both be taken into consideration when determining a level of evidence.

Another issue that has become apparent is that we should be taking baseline risk into account. In other words, a recommendation that may be appropriate for one patient group with a particular risk may not be appropriate for a patient group with a different risk. If you have a treatment that has a relative risk reduction of 25 percent and the baseline risk in the population to which you apply that intervention is 10 percent, you would have to treat 40 people to prevent an adverse event. With that same relative risk reduction of 25 percent but a baseline risk of only 1 percent, you would have to treat 400 people to prevent an adverse event. Your recommendation might be different for those two populations.

Another change is the advent of scientific overviews and meta-analyses. I do not believe anybody should make a recommendation or present levels of evidence unless a high quality scientific overview or meta-analysis has been undertaken.

The first attempt by the Canadian Task Force in 1979 to look at levels of evidence basically distinguished between randomized controlled trials, case-control or cohort studies, comparisons that were obtained between times or places, and the opinions of experts. This original approach continues to be echoed in some of the classification schemes that Dr. Hudgings presented.

A 1984 revision in the hierarchy by the Canadian Task Force, which was endorsed again in 1993 by the U.S. Preventative Service Task Force, places randomized controlled trials at the top, followed by controlled trials without randomization, cohort or case-control studies, comparisons between times or places, and expert opinion, in that order.

This approach has two major limitations. First, it deals only with individual studies. Instead, however, we should look at combinations of studies, because our recommendations come out of a whole literature. The studies may, for example, have different results. Also, there is a risk of bias in terms of what studies are included. A second major limitation is that there no attempt to deal with results, issues of sample size, or power.

Dave Sackett, also from McMaster University, devised an alternative that begins to address these issues. For the first time, randomized trials are separated into two groups—one composed of large randomized trials with clearcut results, and the second containing small randomized trials with uncertain results. The issue of power or sample size comes into play here. Then comes the rest of the hierarchy—non-randomized concurrent cohorts, followed by non-randomized historical controls, followed by case series—once again based on the issues of study design.

The Sackett hierarchy still focuses on a single study. Further, although the issue of results begins to be accounted for, clear separation of magnitude, precision of effects, clinical importance, and statistical significance has not been undertaken.

We developed another iteration. On the right is the Sackett classification, and on the left is an attempt to bring in meta-analyses. Level 1 here would refer to a single randomized trial with low-false positive and low-false negative errors. With a high-quality overview, you could have stronger evidence of from one to one plus. Determining whether an overview might increase a recommendation from one to one plus, or one minus, involves the issue of consistency of results across trials that Dr. Hudgings mentioned. If the trials appear to be homogenous (that is, estimating more or less the same treatment effect), that strengthens the evidence. On the other hand, if the study results are heterogenous (different studies apparently addressing the same or a similar question give different results), the strength of the inference is decreased. This introduces the issue of meta-analysis rating according to the consistency of results.

It is also important to account for the magnitude of the effect. The issue arises when the lower limit of the confidence interval for the effect of treatment exceeds the clinically significant benefit. Basically, suppose you have a big, important treatment effect that you are sure of.

The second class down would correspond to randomized trials with high false-positive and high false-negative errors. Here is when the lower limit of the confidence interval for the effect of treatment fell below the clinically significant benefit. In other words, it appears you may have an effect, but you are not certain. Again, it would be stronger if the results are homogenous, and less strong if the results are heterogenous.

Again, this approach has limitations. There is still confusion between the quality studies and results. It ignores overviews or meta-analyses of lower quality studies, and fails to consider the issue of baseline risk.

I will conclude by describing an approach that we are exploring now, which tries to address these deficiencies. This model assumes that a high-quality overview is available. By high quality we mean that it has addressed a focused clinical question, had appropriate inclusion criteria, and conducted a comprehensive search, where the validity of the studies that have been included are appraised, and where the whole process is reproducible.

Another criterion is that we are able to specify what we might call a threshold number needed to treat, above which the downsides of therapy outweigh the positive effects, and below which the benefits of therapy outweigh the downsides. For example, a theoretical treatment that had to treat more than 100 subjects to prevent an adverse event might not be worthwhile. Treating less than 100 subjects to prevent an adverse effect, however, might be worthwhile.

We try then to separate the methodologic quality from the results. The methodologic quality that we are suggesting is meta-analyses of randomized trials with no significant heterogeneity. The next level down contains randomized trials which do show variability in study results. The next two levels are cohort or case control and noncontemporaneous control. By going from A to B to Y to Z, (instead of C and D) in labeling these levels, we are saying that we see a big drop in terms of the strength of inference between situations in which you have RCTs and situations in which you do not.

The second element in classifying levels of evidence involves focusing on the results and determining whether you have clinically important effects. Assume you have a treatment where you believe an absolute risk reduction of 1 percent is important. That could be a reduction in incidence of adverse events from 2 percent to 1 percent, or from 50 percent to 49 percent. Here are the results of a meta-analysis in which the point estimate—what was observed—would be the highest point on this curve. What we see is that the 95 percent confidence interval around the point estimate is greater than that 1 percent absolute risk reduction. We can then be very confident that the treatment effect is an important treatment effect.

That contrasts with a situation in which the treatment is apparently still beneficial (zero is unlikely to be the treatment effect), but the confidence interval goes below 1 percent. In other words, the true effect of treatment may be less than the 1 percent that we think is the minimally important difference, and our confidence in a treatment recommendation based on these data would decrease.

Similarly, for what we might say are negative treatment recommendations, here is a situation in which the confidence interval around the treatment effect is almost totally below 1 percent. Here, then, we can basically exclude a clinically important effect.

Finally, we have a negative meta-analysis, where the treatment effect is below the zero point, but the confidence interval overlaps the clinically important difference. In this situation we have not definitively excl.. .d a clinically important difference, and a recommendation against using a treatment could not be as strong.

We can then combine this with the concept of baseline risk. Again, with any given risk reduction, the impact or number needed to treat is going to differ depending on your baseline risk. Assume arbitrarily that we have 100 people as the threshold numbe- needed to treat (NNT). If the NNT is less than 100, we think that the therapy is basically worth . If we have to treat less than 100 people, the damage we are doing is not so great to make us hesitate to give the therapy. If the NNT is greater than 100—if we have to treat 200 or 300 people to prevent an adverse event—we will decide it is no longer worth it.

Assume a high baseline risk of about 50 percent, with a 10 percent risk reduction. Here we would need to treat only five people to prevent an adverse event. The confidence interval is likely to be below that threshold NNT, and we would have a confident treatment recommendation. Should the baseline risk drop to 20 percent with that same 10 percent risk reduction, we would have to treat 50 people to prevent an adverse event. Here, our point estimate is still below the line, but our confidence interval is overlapping, and we are no longer so confident that the treatment recommendation should be instituted in that population with a lower baseline risk. If the baseline risk drops further, say to where we might have to treat 200 people to prevent an event, we would recommend not giving the treatment, but if the confidence interval overlaps that threshold NNT, we are not as confident. Or we might have a situation with a very low baseline risk and the same 10 percent risk reduction, where we are confident that we are not going to give the treatment.

All this can then be summarized into Level A evidence, which is from a combination of meta-analyses with no significant heterogeneity, where the confidence interval for A1 is all on one side of the threshold NNT, and A2. where it overlaps the threshold NNT.

For Level B we have F with heterogeneity (different results from different studies) with confidence intervals for . . stronger evidence all on one side of our threshold number needed to treat, and overlapping for the lower level. You can do the same sort of thing for Level Y, which is contemporaneous cohort or case-control studies, and then noncontemporaneous studies for the lowest level.

In summary, then, levels of evidence must be addressed by meta-analyses which may be homogenous or heterogenous. We must take into account both the study design and the results in making treatment recommendations.

DISCUSSION

DR. GRAY: I am the Director of Public Health from Oxford. I am speaking this time not as someone doing a registry but as someone on the board of a health authority that spends a billion pounds a year. Maybe there are two quite different customers for registries. The research managers and research funders are interested in ongoing trials and proposed trials, while people like me want the completed trials. I think when we look for funding for this, it may be useful to keep in mind these two different constituencies, one of which is really interested in published information—in what are we going to do next year for our population—and the other started to think what follow-up research is needed, where you would want to keep in touch with ongoing trials and trials not yet completed.

DR. T. CHALMERS: I hope it was apparent to others as to me that Dr. Hudgings' and Dr. Guyatt's presentations are worlds apart. I was greatly disturbed by one of Dr. Hudgings' statements, because it confirmed my impression from working closely with two of the guideline panels, that if one were restricted to making recommendations based on reliable evidence, the reports would be very short. It is almost as if the fewer the data, the more words. I would like to have seen a guideline saying, "There are no data, fellows. Go out and gather it," instead of, "In the absence of data, here is the way you should treat people."

I have a question for Dr. Guyatt. The word "heterogeneity" has routinely been applied to variations in the size of the difference between the individual studies included in the meta-analysis. His presentation has pointed out that the critical issue could well be the difference in the baseline risk. Although the difference in the result quite often reflects the difference in the baseline risk, there are exceptions, and I think we have to keep straight when we talk about heterogeneity whether we are talking about different patients or different responses.

DR. GUYATT: The heterogeneity that I was referring to is different results in apparently the same populations, with apparently the same interventions. Clearly, if we can separate it according to population, we should have two different meta-analyses: one for one population and one for another. The problem that I think reduces the strength of one's inference is when you apparently have the same population and the same intervention and get different effects. Clearly something is going on that you don't understand, and that to me reduces the strength of inference that comes out of it.

DR. TEMPLE: Dr. Guyatt's presentation really suggests that clinical trials, or at least meta-analyses of clinical trials, should start with a different null hypothesis. Instead of having null hypothesis no effect, the null hypothesis really turns out to be at least an effect of a certain size. That is the equivalent of lowering the confidence interval.

My crucial question is, how do you decide that, and who does it, because everything seems to turn on this. It strikes me as if that is inextricably bound up with the number needed to treat question. One might, for example, say that any reduction in mortality, in other words, null hypothesis of zero, is reasonable when you are talking about death, but whether that is true or not depends on the risk in the population. So those two things seem to be the same, yet you have treated them as different.

DR. GUYATT: We have to, I think, move away from null hypotheses altogether. In other words, we are not trying to say does it work or does it not work, but what is the magnitude of the effect, and what is the plausible range around that magnitude of effect? Because we have to take baseline risk into account, and the same magnitude of effect may be quite appropriate for a population at high risk and inappropriate for a population at low risk.

In fact, and here is where the judgment comes in, we have to decide or have some idea of our threshold number needed to treat—the number that we would need to treat to prevent an event above which it is no longer worth it. That clearly is a judgmental procedure which takes into account a number of things, including the event we are preventing (whether it is a mortality or myocardial infarction or some less serious event), the adverse effects associated with the treatment, and the expense and administrative burden associated with the intervention.

Having done that, we are in a position to decide what recommendation should be made for populations with different baseline risk.

DR. TEMPLE: My problem is that you have made it look as if there are two decisions, and I think there is only one. Take the minimum acceptable effect size, for example. You would not even determine the number needed to treat until you have at least had an effect of that size. I would have said that that does not matter particularly—and that you can accept any effect size as reliable—if, when you then go to the number needed to treat and take your lower bound of what the effect might be, you reach a favorable conclusion. In other words, a very small effect in a high-risk population might meet your test. It is one decision, the number needed to treat, that is at the heart of it. You don't have to make two.

DR. GUYATT: I obviously have not been as clear as I would have liked, because I agree with that 100%, and that is central to what I am saying.

DR. HORTON: A question for Dr. Guyatt again. I understand the importance of baseline risk from your talk, but it is not quite as simple as that, is it? Because how do you judge a baseline risk in a clinical trial when you have a population where the individuals have different baseline risks?

DR. GUYATT: The baseline risk is not judged from a clinical trial. Presumably, one gets the evidence from different types of studies, studies of prognosis, for example, and the patient one has in front of one, as a clinician, is going to come from a different population. So, for example, whether

somebody has a large myocardial infarction and arrhythmias and heart failure, versus a small infarction without those, will create very different risk groups.

DR. HORTON: So the question is, how do you apply population-based evidence, where you have a fixed baseline risk, to the individual patient? How do you make that judgment?

DR. GUYATT: You have to be able to estimate the individual's baseline risk, and that requires appropriate studies of prognosis. In many instances, we have that information. In many other instances, we need more precise information to allow estimation of baseline risk.

DR. HORTON: I am sure most doctors do not even think about baseline risk when they see patients on an individual basis.

DR. GUYATT: Perhaps most doctors do not think of randomized trials either, but we are trying to change that.

DR. BOISSEL: Jean-Pierre Boissel from Lyon, France. If I may make two comments. Instead of speaking about heterogeneity, why not speak about either residual heterogeneity, or unexplained heterogeneity?

DR. GUYATT: In fact, I do mean residual heterogeneity after you separate out population and intervention issues.

DR. BOISSEL: There are other ways to deal with heterogeneity than to split a population in two or three parts.

DR. GUYATT: Really I am talking about unexplained heterogeneity. You are quite right.

DR. BOISSEL: My last comment is about basic risk. The risk of the control patients is an estimate of this basic risk.

DR. GUYATT: Yes, except that, even within a control population, you have some high-risk people and some low-risk people. Ideally, the clinician will not just assume that everybody is in the middle, but will take into consideration whether her patient is a high- or low-risk patient.

DR. SIMES: I have two comments concerning classifying the strength of evidence according to meeting a threshold or not. First, I think the number needed to treat is a useful concept, but it does not cover all ways of saying whether a size of effect is adequate or not. When you have relative reductions over time, for example in some cancer studies, it is not necessarily the appropriate way, but I think some measure of saying whether the size of the effect is above or below a threshold is useful.

My concern is that we use that as a way of classifying strength of evidence, as opposed to simply classifying studies as to whether the evidence is in one case or not. Suppose, for example, you had two meta-analyses of the same question—one that demonstrates that you reached the threshold; another, looking at the studies in a different way, which did not. I would not like to say one of those was, in fact, better evidence than the other.

DR. TEMPLE: It strikes me that the usual problem about interventions is not whether they work enough, but whether they work at all; and that our major quest is to try to get people to do randomized trials to get their information, instead of trying to deduce things from databases that are not good enough.

I am worried about the idea that, having shown an effect—that is, a difference from zero—the challenge then becomes to show exactly how big the effect is. This means sample sizes have to grow enormously, I am a little frightened about discouraging the process when people realize that, instead of the sample sizes of 20,000 and 30,000 that we are now accustomed to, you have to go to 40,000, 50,000 and 100,000.

(This concluded the discussion.)

ROLES AND USES OF CLINICAL TRIALS REGISTRIES— MEDICAL DECISIONS, RESEARCH, NATIONAL AND INTERNATIONAL COOPERATION I

T. Chalmers

DR. T. CHALMERS: The program announced that I was from Harvard and the VA, but I am no longer at either, except for a token teaching appointment at Harvard. I am now at Tufts and Dartmouth.

I want to make two points. One is that the registry of clinical trials should include a registry of meta-analyses of clinical trials, because they are becoming as hard to keep track of as the original research. Second, if we had had a registry of clinical trials for the last 20 years, I believe we would have saved many lives.

A recent tabulation of the number of meta-analyses of randomized controlled trials which we have in our files in hard copy and on diskette demonstrated a rapid rise which shows no signs of leveling off. We find that we now have the same problem with meta-analyses that we had before with clinical trials—we need a better method of keeping track of what was done in the past, especially since meta-analysis was not an indexing term until quite recently.

To address the problem of how lives could have been saved if we had had a registry for the last 10 to 20 years, I must explain cumulative meta-analysis. The 38 trials of intravenous streptokinase therapy in acute myocardial infarction were compiled. The usual meta-analytic method of presenting the data when the odds ratio is the scale was used. A cumulative meta-analysis, defined as doing a new meta-analysis each time a new trial appears was done. A mixed results, including studies in which mortality is significantly reduced and others in which mortality is increased by streptokinase but not significantly, was seen. The meta-analytic result showed a highly statistically significant reduction in mortality, which added up to appproximately 30 percent.

It was shown that by 1973 the cumulative meta-analyses was statistically significant. By 1988, and we now have examples in which untreated controls were used going through 1990, it was absolutely clear that the death rate was reduced by streptokinase. I do not want to get into the argument of when you make the decision that enough evidence is accumulated. However, where GISSI-1 comes in in 1986, its cumulative effect was imperceptible; and when ISIS-2 is added in 1988, it did reduce the confidence intervals a little. However, the effect was well-established long before those trials were thought of or completed.

This is the same data in regard to beta blockers after myocardial infarction. A few more significant studies were ongoing, reflected by the fact that the cumulative meta-analysis was positive

by 1980. BHAT became famous in 1982 after it was stopped because the number of patients in the study was sufficient to indicate, in that study alone, that a statistically significant result. This was considered a landmark case, because the result was reached before the projected number of patients. What should have been noted, however, was that the same answer would have been available 4 or 5 years sooner if meta-analysis of randomized trials had begun. That would have required a registry to find them, because finding all trials has been a major effort.

The next step in this argument lies in the reaction of the community to the randomized controlled trials which were put together in meta-analyses between 1982 and 1990. We have categorized that reaction to be routine administration of the drug, to be used only in specific patients, rare or never, still experimental, or not mentioned. Note that although this highly statistically significant effect—and this now is cumulative by year instead of by study—was clearly apparent between 1973 and 1978, people did not say that streptokinase should be routinely administered to patients with acute myocardial infarction unless they have a contraindication until after GISSI and ISIS.

This contrasts with the data on intravenous Lidocaine. In this case, "routine" means every patient with an infarction, and "specific" means for all who have ventricular extrasystoles. Notice that the experts were recommending Lidocaine throughout this period when the randomized controlled trials showed that it had no effect and the tendency was toward increased mortality. In the one meta-analysis that we did, in which we segregated the patients into two-day mortality versus later (i.e., while the infusion was going in the death rate from intravenous Lidocaine was statistically significantly higher in the case of Lidocaine than in the case of the placebo over the first 48 hours. That washed out when you followed through the whole length of the study. Lidocaine is not being used much anymore, but at the time the trials were somehow lost sight of by these experts who wrote review articles saying the drug should be used in just about everybody.

In the case of long-term beta blockers, there still was a 5- to 8-year lag between the motile recommendation that the drug be used. This may be due to the fact that it was hard for these people to keep track of the randomized controlled trials which they should have been reviewing.

This is in another field, the effect of endoscopic therapy on the emergency surgery rate. I have been keeping track of the emergency surgical rate in GI hemorrhage and the death rate for about 30 years and, although the surgical intervention rate has risen precipitously, the death rate has stayed exactly the same for 30 years.

In this case the scale is different because we are using the random effects model of DerSimonian and Laird, and giving the risk difference on one side and the control rate risk on the other side, and some of the data that showed that they are reasonably correlated. The cumulative meta-analysis shows again that by 1984 the therapy had been proven to be significantly reducing emergency surgical

intervention, that the surgical intervention rate stayed the same throughout these years, and that the effect stayed highly significant.

Note that the overall death rate, which had not changed in 30 years, is reduced by about 30 percent with the advent of endoscopic therapy—reduced 3 percent absolute from some 10 percent, that it had been. Notice also that the control death rate is staying the same here at about 10 percent. But the efficacy of therapy is highly significant, and that might have been revealed by a registry of clinical trials.

In this example, the meta-analysis was reported in 1981, demonstrated that antibacterial therapy around the time of colorectal surgery reduced the wound infection rate significantly. By the time the third study came out it in 1971, it was statistically significant.

What were the investigators thinking about in 1979 and 1980 when they decided that they should randomly assign half the patients undergoing colorectal surgery to a placebo rather than an antibiotic? It is possible that they were unaware of the other studies, which having a registry would address. It is also possible that they did not believe meta-analysis, and that is changing. They also might have thought that there was no need to prevent the infection because it could be treated later.

The cumulative meta-analyses for peri-operative deaths showed that by 1977 antibiotics were reducing peri-operative mortality. Thus, it is not only a matter of treating the disease rather than preventing it, because the patient would not be around to be treated.

The story of hysterectomy is one more example. The rate of total infection or febrile episodes was determined by the cumulative meta-analysis fixed effects model of antibiotics for hysterectomy. I did not bring the data dichotomizing this into abdominal and vaginal. We found no difference in the randomized trials between vaginal and abdominal. They both were dramatically effective. We could not gather data on mortality in hysterectomy, because the mortality rate is so low that it would have taken more patients than the 3,490 in the study, with mortality well under 1 percent. You can see that the cumulative meta-analysis indicates that antibiotics are highly effective, and somebody made a mistake in randomly assigning patients to a control group and failing to use antibiotics prophylactically when doing a hysterectomy.

In sum, it is imperative that we develop a database of randomized controlled trials to which practicing doctors can refer. Then if they find there is no evidence in favor of one treatment or the other, they can randomize, rather than mistreat patients on inadequate evidence.

ROLES AND USES OF CLINICAL TRIALS REGISTRIES— MEDICAL DECISIONS, RESEARCH, NATIONAL AND INTERNATIONAL COOPERATION II

Iain Chalmers
Cochrane Collaboration
Oxford, England

DR. I. CHALMERS: A curious double standard is applied throughout the biomedical research community, not just in the trials community. It is illogical and, as Tom Chalmers has already illustrated, dangerous. The community usually demands that scientific principles be observed when people conduct primary research, and usually acquiesces when scientific principles are ignored by people conducting secondary research or reviews. It is a completely unsustainable double standard. Good primary researchers are expected to write protocols for their primary studies and to conduct them according to scientific principles. When these same people are asked to write a review chapter for a book or for an article in a journal, however, they basically ignore those principles.

In secondary research or reviews, as in primary research, steps must be taken to minimize biases, systematic errors, and imprecision. A systematic review should start by stating the objectives of the review and the eligibility criteria for data to be admitted to the review. The next stage is to identify potentially eligible studies, and this is where we currently have difficulties. The scientific reviewer will then apply the prestated eligibility criteria to the identified studies, and then assemble the most complete dataset feasible, possibly with the active participation of people involved in the primary studies. Then an analysis will be conducted and, sometimes, a statistical synthesis or meta-analysis. (That, however, applies more to reducing imprecision than to reducing bias in reviews.) Then, a structured report will be produced describing what has been done. Not everyone who reads the report will agree with those methods, but at least they will know what has been done. Then if they want to replicate the research using different methods or different eligibility criteria, they can do that and report their findings.

The concept of a double standard is actually very recent within the biomedical field. It has been far more prevalent in psychology and especially educational psychology. I want to illustrate that by quoting from an editorial written by Dr. Relman in 1980. This was, in part, a comment on the consensus development program which the Office of Medical Applications Research is responsible for. He said, "The NIH consensus development program is certainly a useful exercise, but it has its limitations. Assessments of current knowledge, no matter how sophisticated and rigorous, cannot go beyond that knowledge and rarely generate any new information. What was unknown before the

review remains unknown afterwards.. ." He then calls for more primary research by continuing". The best way to assess clinical practices is to generate new information."

I do not believe that the issue of sophistication and rigor in reviews had even crossed his mind in any serious way. The evidence to support that suggestion comes from the first of what I would say are two of the most important analyses ever published in medical journals. The first, published in the *Annals of Internal Medicine* and welcomed in an editorial by Ed Huth, was an analysis done by Cynthia Mulrow, currently at the University of Texas. She analyzed the scientific quality of review articles published in the *New England Journal of Medicine*, the *Annals of Internal Medicine*, the *BMJ*, and the *Lancet*, and stated, "Current medical reviews do not routinely use scientific methods to identify, assess, and synthesize information."

You have already seen some of the results from the second of the two studies that I am commending. Tom Chalmers has written many important papers in his time, and the conclusion that he and his colleagues came to was, "Because reviewers have not used scientific methods, advice on some lifesaving therapies has been delayed for more than a decade, while other treatments have been recommended long after controlled research has shown them to be harmful."

Most of the textbooks that were studied in that particular analysis were American textbooks, but I want to make it clear that the problem is not confined to this country. The *Oxford Textbook of Medicine*, Second Edition, 1987 states, "The clinical benefits of thrombolysis, whether expressed as improved patient survival or preservation of left ventricular function, remain to be established." It is lethal advice because of the double standard that is being applied by the biomedical community.

Two years before that was published (and this is only one of a number of systematic reviews of thrombolysis and acute myocardial infarction, some published earlier), people who had done a systematic review of the randomized trials came to this conclusion. Because all of these trials were small, their separate results appear contradictory and unreliable, but an overview of the data indicates that the treatment produces a highly significant reduction in the odds of death, an even larger reduction in the odds of reinfarction, and an absolute frequency of serious adverse effects to set against this that is much smaller than the absolute mortality reduction.

People talk about the applicability of the results of randomized controlled trials in general medical practice. As someone who is at risk of a myocardial infarction, I wish to receive thrombolytics if I have a myocardial infarction. I do not wish my treatment to be decided on some database analysis which has not drawn on the results of randomized controlled trials.

Archie Cochrane wrote in 1979, "It is surely a great criticism of our profession that we have not organized a critical summary, by specialty or subspecialty, adapted periodically, of all relevant randomized controlled trials." He was not someone who was saying that there were not forms of care

whose effects could be identified confidently without randomized controlled trials. No randomized trials were needed to identify thalidomide as the cause of phocomelia; the effect was very dramatic. No randomized controlled trials are needed to know that, if you replace a hip with bad osteoarthritis, you relieve people's immobility and pain. However, for most things that are done within the health services, you need to control biases by randomization. Otherwise you may be misled into thinking that a treatment is effective when it is not, or vice versa.

In response to Archie Cochrane's challenge, a pilot study began in 1978 and is ongoing. It is a systematic review of RCTs of care in pregnancy and childbirth. Data collection for systematic reviews should begin by identifying as high a proportion as possible of the methodologically trustworthy studies. Again, I stress that this is not a rule that applies just to randomized trials. It applies to all activities that require a systematic look at what evidence has been generated from empirical research.

If data collection is to begin that way, Phase I of data collection for looking at randomized controlled trials and their results in pregnancy and childbirth is to identify the trials. Currently, this is monstrously difficult.

We spent about 7 years trying to identify trials thoroughly. We searched MEDLINE back to 1966 using the best search strategy we could devise. We also hand-searched 70 journals back to 1950. To try and flush out unpublished trials, we did a mail survey of over 40,000 clinicians in 18 countries. We came up with a registry of about 4,000 trials, and that is being maintained at the moment. I want to emphasize the enormous amount of effort that needs to go into actually identifying the trials in the first place, because that is, indeed, what part of this conference is about.

Phase II was to use that trials registry for preparing systematic reviews between 1985 and 1988, and for keeping those reviews up to date, as new evidence becomes available, forever. We owe that to patients. That is what the Cochrane Collaboration is trying to do. It is people all over the world who are trying to prepare, maintain, and disseminate systematic reviews of the effects of health care.

In conclusion, clinical trials registries are essential for efficient data collection for systematic reviews of the effects of health care. They are essential because those reviews are something we owe the public, which has invested and participated in the research that has been done.

Every country can and should help to establish trials registries. It is interesting that 30 percent of the women represented in the early breast cancer trial collaboration analyses are Japanese, and that before 1985 none of the breast cancer participants in the West had any idea that there was that much randomized evidence available in Japan. That is an example of the sort of "gold mine" that is out there waiting to be discovered by a more organized and systematic approach to trials registries.

I believe that the United States is uniquely placed to coordinate the global effort needed to establish and maintain the clinical trials registries that are required, because it reports and does more randomized controlled trials than any other country in the world. Also, with the National Library of Medicine, it has by far the best bibliography of controlled trials currently available.

Let me illustrate what is meant by keeping a review up to date and the commitment that entails. I will use diethylstilbestrol, DES, for this example. Diethylstilbestrol was a drug which was introduced in the late 1940s in the belief that it helped women avoid miscarriage and stillbirth. It was introduced on the basis of uncontrolled trials of database analyses, historical controlled studies, and so on. In fact, five reasonably well-controlled trials were done between 1950 and 1955, but there was no way of easily distinguishing those trials from the other body of evidence that was being looked at.

However, in retrospect, we can see that there are six trials represented in this particular overview. These are the outcomes that have been studied in one or more of those six trials. You can see that there was quite confident evidence from those controlled trials that this was not a sensible treatment to be offering women at all, outside the context of controlled trials. In fact, the likelihood of no surviving infant was actually increased when you looked at the results of those five trials.

Now, the five trials obviously would not be on MEDLINE, because MEDLINE started in 1966. But as you look at the outcomes, you notice that both the mothers who had been given the drug, and their children—boys and girls—suffered as a consequence. That information comes from 30 year follow-ups of three of the randomized controlled trials.

The data for miscarriage indicate that the actual trials contributed to that result. Amongst them is a randomized trial reported in 1977. Because Harrup reported the adenocarcinoma side effect of this drug to the *Journal of the American Medical Association* in 1969, how could it be that people were still being randomized into trials of DES such that they could be reported in 1977?

I think there are a number of lessons from this. That particular trial was in Geburtshilfe Frauenheilkunde. We discovered it through MEDLINE, without journal-searching. Unless people do systematic reviews of the randomized controlled trials, not only will they be giving ill-informed treatment to people, but they also will be doing unnecessary research, a point that Tom Chalmers made earlier. Trials registries are the starting point for this whole process. Unless they are established, this cannot be done efficiently. Unless this is done efficiently, research-funding organizations will make wrong decisions about which trials to fund.

I leave you again with a plea that we really take seriously, as an ethical issue, the challenge to actually develop these registries more satisfactorily.

ROLES AND USES OF CLINICAL TRIALS REGISTRIES— MEDICAL DECISIONS, RESEARCH, NATIONAL AND INTERNATIONAL COOPERATION III

Kay Dickersin

DR. DICKERSIN: I am going to review the rationale for maintaining registries and databases. Again, the types of registries we are talking about include completed randomized trials and other studies, both published and unpublished and, if possible, registries of ongoing trials.

First, why do we create big databases, let alone registries of trials? We want to gather and control a large amount of information. GenBank, a database of nucleotide sequences, is an example. There is a similar database for amino acid sequences of proteins. Molecular biologists saw the reason to do this long ago, and I think we should, too. The second rationale is to give notice that something exists. This is very important.

The last reason is to prevent duplication of effort. For example, both the cataract PORT and the guidelines on cataracts are working on very similar literature reviews. Because they are funded by the same agency, both groups know that the other exists and have been able to work together to prevent duplication of effort.

The same reasons apply for trials registries. Again we want to gather and control a large amount of information, though in this case not the results of trials, but the key information about their design and operation. We also want to give notice that trials exist, both those that have been completed, and those that are ongoing. Further, we want to prevent duplication of effort. There are other reasons including to help recruitment to ongoing trials, to facilitate systematic reviews, and to prevent publication bias. Registries could also help enhance collaboration. Finally, registries can enable methodologic research. We have been able to use registries to look at the quality of randomized trials that are being done, and I will provide examples of those.

There are many reasons why different groups establish registries. The NIH and other funding agencies have had inventories of trials because they want to know what they are funding. For example, from about 1975 to 1979, the NIH had an inventory of clinical trials. The OMAR registry of clinical studies still exists and serves in part as an inventory of ongoing human studies at the NIH.

Spain started an inventory of randomized trials in order to centralize ethics review of the country's studies, since all studies are essentially funded by the government. They wanted to know what was ongoing, whether local ethics committees had approved these studies, and whether they agreed with those decisions. The result of that inventory is that they have a registry of clinical trials that is mandated by law.

The NEI puts out a booklet of all the trials that it funds. I believe that the purpose of this registry is to serve as general information for investigators and others who might want to know about the clinical trials NEI is funding.

Databases such as Physician Data Query (PDQ), sponsored by the National Cancer Institute, and the various AIDS registries serve primarily as information sources to patients and providers, rather than to researchers. With PDQ, for example, providers or patients who know how to use MEDLINE or an online database can learn about relevant ongoing clinical trials in cancer. Finally, groups such as the Oxford group have used registries as a basis for systematic reviews, as Iain Chalmers explained about the childbirth database.

At the University of Maryland we have been keeping a registry of registries for a number of years. We are currently sending out questionnaires to all known registries to find out if they still exist, how they do their updating, and how many trials they now include. This slide shows the data we had as of a year ago. All the registries shown cover AIDS, and most are regional registries within the United States. Their main purpose is to let AIDS patients know that trials exist.

Most of these are funded by national or regional government. There also is the AIDSTRIAL database, which is federally sponsored. I think these show how successful patients have been in getting registries started in one field. Note that most of the registries I am showing are of oning, not published trials.

This second page of registries is more diverse. The first four cover cancer and include PDQ, the NCI cancer control intervention studies, and CRISP, sponsored by the Public Health Service. As with most of these, it contains more than just trials. There also is a registry for dentistry, the VA has a registry of its studies, and so forth.

Some of these registries are internal and not meant for general use. As I said before, the reasons differ for having registries.

I want to give some additional examples of how trials registries can be useful for research. The NIH 1979 clinical trials inventory funded a study that yielded two types of information. We followed all trials that were funded by the NIH in 1979, and part of their inventory, and were able to learn whether the trials were published and their results. We also looked at whether the results were related to publication. In addition, we were able to say whether women had been included in trials funded by the NIH at that time.

Another example is the Oxford Database of Perinatal Trials, a pilot study that has evolved into two databases that are mainly for internal use so that systematic reviews can be performed. We have used that database for a study of MEDLINE searching and have also looked at it to see whether abstracts are published in full and at what rate.

This is the study of publication bias where we found that, of trials funded by the NIH in 1979, there were 74 with results that we classified as nonsignificant, and 124 with results that we classified as significant. You can see that 85% were published in the nonsignificant group, and 98% in the significant group. This study is unusual in the high percentage of studies that were actually published, but there was a strong association between significant results and publication.

These results were from the same study. We had 293 trials that were registered on the NIH inventory and found that 13 excluded males and 12 excluded females. Of those that excluded males, all were relevant only to females: studies of vaginitis, childbirth, and so forth. Of those that excluded females, however, only four were relevant only to males. There were two we could not classify. Both were drug trials, and women may have been excluded because of the possibility of pregnancy, although there was nothing on the database that said this. Five of the trials were relevant to both sexes. These were the cardiovascular trials that have been highly publicized in the media.

Using the Oxford Database of Perinatal Trials, we followed abstracts to see what percentage reached full publication and whether or not this was related to the results. You can see that there were 98 abstracts that had positive results, and 78 abstracts we classified as having neutral or negative results. The publication rate was in fact higher for the results that were neutral or negative, indicating no publication bias from this particular study.

Finally, we used the Oxford Database to look at the sensitivity of MEDLINE searching. As you can see, there have been three studies so far that have used the Oxford Database to compare a MEDLINE search with the contents of the Oxford Database. The range of studies and sensitivities that were found were presented.

HEALTH CARE REFORM

William Harlan

DR. HARLAN: First, it is clear from the different legislation that has been introduced that the intent of health care reform is to provide reasonable reimbursement only for efficacious treatment, implying comparison of costs across different treatments that might be similar in efficacy. The guideposts to most of this information, I think, would be randomized clinical trials. However, they are not the only source of information, and we need to keep that in mind as we talk about registries.

Another likely result of health care reform in this country is an increase in the number of clinical trials being done, and an increase in the amount of health care research being done. I think this is unavoidable, because that is the basis on which we would like to make our judgements about reimbursement. An ongoing clinical trials registry would help in several regards.

First, it would help in terms of the issue of redundancy of trials. In that regard, I think it is extraordinarily important that any ongoing clinical trial registry be international in scope.

There is a part of the Health Care Reform Act introduced by the Clinton administration that may have escaped your notice, but is very important. It says that the costs of standard medical care for a condition being investigated in an approved trial would be paid for by the health plan. This means that if you have hypertension for example, the usual costs of taking care of your hypertension would be paid along with the test for monitoring of care if you are in the clinical trial, and it would not be charged to the clinical trial. I think this will further the support of clinical trials, but it also heightens the need for a registry where we can find so-called "approved" clinical trials; those that have been approved by NIH, PHS, the Department of Defense, Food and Drug Administration, and the Department of Veterans' Affairs.

Also compelling to me is a registry of ongoing clinical trials. Looking at clinical trials over the years, we have moved from a very strong base within the academic arena to the large streamlined trials—such as ISIS, GISSI, and the digitalis trial—where we acquire patients in a large number of clinical centers and where patients may be accrued in the practice setting. The information gathered thus becomes more immediately relevant. It also trains the practitioners to use the results of those trials in their own practice. It is a wonderful educational program.

An ongoing clinical trials registry providing information about the trial and how to enroll as a patient or provider would also reduce trial costs. Further, if we have a registry of trials, it will allow those interested in economic analysis and the like to identify the trials where they might collect information that would improve our knowledge of the clinical trial itself.

If we go to an alliance kind of arrangement for clinical trials, perhaps the alliances might think of themselves as a clinical trial group. An alliance formed within Washington, D.C., for example, might think of itself as needing to do particular clinical trials and could coordinate with alliances around the country with similar interests.

I question whether we want to require randomization among all trials included in the registry, because randomized trials do not always contain the information on which decisions are based. Should we include nonrandomized trials as well? Also, who would participate in such an ongoing registry? I believe it should be international, and that it should include all groups doing clinical trials.

This has implications for the amount of information that should be included in the registry. What is the minimal database that would allow for identification and participation in the ways that have been discussed today? Also, what is to be left out of the registry (and to be obtained by contacting those conducting the trial)? I would argue for a minimal database that is most informative to all potential participants rather than a very complete database, in order to minimize the specificity and detail and to maximize the ability to identify all of the studies that are underway (sensitivity).

These questions apply for ongoing clinical trials registries, for which we heard arguments this morning. In fact, a registry of completed clinical trials can come from existing infrastructure and databases, with some improvements. I think we also need an ongoing registry of clinical trials, and I submit that the particular interests of health care reform in this country form a good basis for the kinds of information and studies that ought to be included.

WOMEN'S HEALTH AND CONGRESS

Judith H. LaRosa, Ph.D.
Office of Research on Women's Health, NIH

DR. LAROSA: I am here to talk about Congress and women's health. The storm to create equity in research for both genders has been gathering for decades. Female and male scientists, clinicians, women's health advocates, women, and their families were tired of not knowing what to do when it came to women's health issues. The storm broke when the General Accounting Office, at the request of the Congressional Caucus for Women's Issues, released its report showing gender inequity in biomedical research supported by NIH.

In its wake came numerous actions. One was the creation of the Office of Research on Women's Health at the National Institutes of Health. Its mandate is to establish a research agenda on women's health and to assure that women are appropriately included in studies of all diseases, disorders, and conditions that affect them—not just those traditionally thought of as women's health issues (i.e., diseases of the reproductive organs).

The storm has lessened only slightly in its intensity. Since 1990 when the Office was established, action on women's health has increased throughout the health-related agencies of the federal government. Most agencies have an office with resources for action, and I think that is very important. These are not just window-dressing offices. Not only are the agencies expending resources on women's health, but representatives of these agencies work together formally on the Public Health Service Task Force on Women's Issues, as well as informally.

In the three years since the Office of Research on Women's Health was established, women's health has begun to gain a permanent and more equitable place in biomedical research. Congress, however, is not content to hope that it will become permanent and has taken firm steps to mandate such permanency.

These steps are perhaps best defined in the NIH Revitalization Act of 1993, which became public law on June 10, 1993. The law gave many of the NIH institutes and centers specific direction for women's health efforts. For example, research on breast and gynecological cancers was assigned to the National Cancer Institute; a program of research on osteoporosis to the National Institute of Arthritis, Musculoskeletal and Skin Diseases; aging processes regarding women to the National Institute on Aging; chronic fatigue syndrome to the National Institute of Allergies and Infectious Diseases.

The Office of Research on Women's Health was mandated with that legislation, along with an advisory committee and tasks in addition to its research agenda. Efforts undertaken by the Office

included revising the 1990 guidelines on the inclusion of women and minorities in clinical research. Guidelines published as a result of the NIH Revitalization Act are intended to ensure that all future NIH-supported research involving human subjects is designed to elicit information about individuals of both genders and all racial and ethnic groups. Further, clinical trials should examine differential effects on such groups. These guidelines will be published shortly in the *Federal Register*, to be followed by a comment period of approximately one year.

The 1990 guidelines on the inclusion of women and minorities were presented. While our office is tasked specifically for women's health, I include minorities because women are minorities in some cases. Adequate numbers of women must be included in clinical studies (i.e., studies of human subject research) in proportion to their prevalence in the condition under study. If this does not happen, it will affect the priority score.

The Revitalization Act of 1993, keeps the essentials of the 1990 guidelines and expands them. Women and members of minority groups and their subpopulations are to be included in all human subject research. With Phase III clinical trials, women and minorities and their subpopulations must be included in numbers sufficient to permit valid differential analyses. (Note that "valid analysis" does not mean "statistically significant.") Congress further directed that cost cannot be used as a reason for not including these groups in clinical trials.

Also different in the 1993 version is that the NIH has been tasked to work closely with the research community to develop programs and activities of outreach for the recruitment of women and minorities into these studies. Going beyond mere recruitment, the Office of Research on Women's Health has added "and retention."

The intent of these guidelines is to establish a continuum of research, both basic and clinical, that provides scientists considering a Phase III clinical trial with the information they need to decide whether to test for gender, racial, and ethnic differences. We need to obtain these data on minorities and both genders early in the research process, when hypotheses are being formulated, baseline data are being collected, and various measurements for instruments and intervention strategies are being developed. Further, it is imperative to obtain the data from a variety of sources including observational clinical study data; basic laboratory data; and clinical, physiologic, biochemical data from Phase I and II studies, genetics studies.

We are looking for substantial differences. We urge you to read the guidelines closely and to understand that we are trying to tell the scientific community that we understand the concerns. We, too, are scientists. In essence, we are saying that one must be mindful of these issues throughout the study-construction process.

Congress is requiring of our office, in regard to a clinical trial registry, to work in consultation with the Director of the National Library of Medicine to establish a data system "for the collection, storage, analysis, retrieval, and dissemination of information regarding research on women's health that is conducted by the national research institutes." It is important to note also, that the information is to be "available to health care professionals and providers, researchers, and members of the public."

This speaks directly to what we are to do with the clinical trials registry. We feel quite keenly that a clinical trials registry is very important, and we support much of what has been said here today. We do not, however, believe that a clinical trials registry that directs itself wholly to women's health is prudent or feasible. We feel quite keenly that, especially in fiscally conservative times, it is prudent to ensure that research addresses both men and women, age, ethnicity, and so forth. We are developing databases and information that will allow investigators, now and in the future, to make judicious decisions about the design of future clinical trials.

Finally, concerning maintaining and operating a program to provide information on research and prevention activities, I think it is very important to have at least, as Dr. Harlan said, a minimal amount of information that will enable the investigators to go back and look at where they need to find further information and to whom they need to speak.

We believe that the sooner we fully integrate women's health into biomedical research, the sooner it will become a permanent part of the fabric of biomedical science, from which it can never be removed. We want it to become a nonissue, so that when studies (particularly clinical trials) are designed, if the disease, disorder, or condition affects women, women will be included in the hypotheses if appropriate.

RESEARCH INTEGRITY

Ed Huth

Online Journal

DR. HUTH: I will address possible ways that editorial functions and the functions of journals might be improved through access to clinical trials registries. Some of my remarks may apply only to registries of published trials, and some to registries including unpublished trials. I will not distinguish between the two.

First is the chronic problem editors have in maintaining adequate files of truly competent peer reviewers. Most major clinical journals maintain quite comprehensive files of potential reviewers. At the *Annals of Internal Medicine*, people regarded as likely to be competent to serve as peer reviewers are invited to indicate whether they will serve. They are sent a questionnaire. Then they return their terms, indicating their particular interest and competence, and the terms are translated into MeSH terms, providing an established vocabulary.

The problem is that the reviewer's own specification of interest is not very detailed. For example, the reviewer may indicate an interest in the treatment of hypertension, but there is nothing about the particular drugs or class of drugs that he or she may have special competence with. With access to registries of clinical trials that include indexing for drugs of interest, editors might obtain more expert peer review.

The cost of updating files is another problem. Questionnaires must be sent out periodically, and considerable time is spent keyboarding new information into the computer files. The file of potential reviewers kept at *Annals of Internal Medicine* includes approximately 9,000 names, yet various topics are still poorly represented. This provides a sense of the problems involved in maintaining such files.

Related matters are the need to invite editorials and other kinds of synoptic information. This is a subset of the problem of being able to find competent persons for particular editorial needs. Finding methodologic experts is especially difficult. Many potential peer reviewers may have methodologic competence equal to that of persons specializing in study design and analysis, but the number of statisticians and persons with competence in critiquing study design is but a fraction of the total population available for reviewing. Having access to a larger number would certainly be helpful.

The next item is, "Aid in Identifying "Rogue Reports." "Rogue reports" is a term I coined for situations in which a member of a multicenter trial decides to take the data from his or her own institution and publish a separate report. These situations can lead to conflicts with trial collaborators and duplication of data. A trials registry might provide a better chance of identifying such rogue reports prior to publication.

Next is an interesting potential problem. There is some chance that, as electronic journals develop, their content will be "invisible" as compared to paper journals. Also, the number of reports may multiply, and such reports may be less likely to be known to peer reviewers. With access to the registries, reviewers (who might themselves be managers of registered trials) may be better able to identify cases of plagiarism involving material taken from papers published in electronic journals.

Next, would it be possible for journals to publish protocols at the time of their registration? Would there be value in making these available in an electronic journal, for example, for external critiquing? Then, as the trial proceeds, there might be a wider range of critiquing available to the conductors of the trial.

If protocols were maintained through the entire study, would it be possible to publish them before results? Then, when the data are available, subsequent reports would not have to include long methodological sections.

Next is the notion of developing a standardized taxonomy and nomenclature for trial design and data components as a product of the design for the registry. This would provide a standard vocabulary for journal reports, structured abstracts, and indexing. As Dr. Schuyler indicated earlier, substantial steps in this direction have already been taken. Further, this taxonomy and nomenclature will facilitate accurate access in bibliographic and whole-text electronic media.

These conclusions and suggestions are based on some simple assumptions. First, if registries are to function usefully, they should be available online, and to those not directly involved in clinical trials work. Further, the registries' thoroughness of indexing and efficiency of access to information for a variety of possible factors and components will be an important determinant of their utility.

EXPERIENCE OUTSIDE U.S. I

Jean-Pierre Boissel, M.D.
University of Lyon

DR. BOISSEL: I will focus on my experience with the ISTH registry of planned and ongoing clinical trials. As mentioned this morning, there are various reasons for registering planned or ongoing clinical trials. One is to keep track of unreported trials. We are all familiar with the file drawer or iceberg phenomenon, along with publication bias, referred to by Dr. Dickersin. My second reason is to be aware of trials that are currently being done, and this applies to the scientific community, ethical boards, and consumers of health care.

The first objective of the International Society of Thrombosis and Haemostasis and Council of Thrombosis registry is to have a registry of planned and ongoing trials, to help investigators in planning or designing new trials and/or performing meta-analyses. "Performing meta-analyses" was added in 1985. The trials to be registered are multicentered, randomized controlled trials. No definition of randomized controlled trials was specified, and deciding which trials are controlled and randomized was left to the registry keeper. The ends were defined as the field of thrombosis in both studies and atherosclerosis.

During a meeting of the International Committee on Thrombosis and Haemostasis in Basel in September, 1974, a decision was made to start a continuous registry of major trials in progress with haemostatic or antithrombotic drugs. The first report of the subcommittee on clinical investigation, led at that time by Professor Mark Verstraete from Louvain, was presented and, one year later, published in the society's journal.

The present organization of this registry includes a network of potential investigators and trial sponsors, including governmental research institutions and pharmaceutical companies. A management team is located in Lyon. The chairman, a part-time physician, is in charge of maintaining, updating, editing, and disseminating the registry data. A part-time secretary and a part-time programmer are also part of the team. Finally, a board of two auditors, consisting of the Chairman of the Council of Thrombosis and the Secretary of the International Society of Thrombosis and Haemostasis audit each yearly report before it is disseminated.

To facilitate registry management, the clinical trials have been assigned to various categories. Currently, there are ten. We can increase the number of categories whenever a new area is considered by a new clinical trial.

For each registered clinical trial, the registry contains: the acronym and its meaning; the objectives of the trial; how patients are selected; the tested treatments; the control treatment; the

dosage of any drugs involved; the treatments compared; the study design; the randomization, with a special emphasis on the procedure used to randomize patients, in order to be sure that the trial is purely randomized; the outcomes (especially the clinical outcome); central validation procedures; treatment duration; follow-up duration; the number of sites; the sample size; the expected and the actual beginning dates; the expected date of completion; when available, the actual date that recruitment is completed; the name and address of the coordinating officer or the principal investigators; the data center; the audit center; the central laboratory, or laboratories, if any; and the name of the sponsor.

Since the first report in 1974, the total number of clinical trials registered in the ISTH/Council of Thrombosis registry at the beginning of 1993 is 419. The next report should be published shortly.

Other European clinical trial registries include several registries in the United Kingdom, including the ARC database of clinical trials in rheumatoid arthritis, the British Heart Foundation register of cardiovascular trials (which I believe is still in the start-up phase), the Medical Research Council cancer trial office registry of cancer trials, and a clinical trial registry for AIDS treatment. A French registry is kept by the Federation of the Anticancer Center. Dr. Dickersin already mentioned the Spanish governmental database of clinical trials. Also, most European drug regulation agencies have databases of clinical drug trials, although they are not accessible to people outside the agency. Another registry will be funded by the European Union on European AIDS trials.

We have encountered several problems with the ICTH/Council of Thrombosis clinical trial registry. The first problem relates to data collection. We receive contributions from researchers and research agencies. However, we have found it difficult to obtain data from regulatory agencies, particularly in Europe, and from companies. Some big companies, and especially North American companies, will not contribute to the registries. In addition, there are the ethical boards, which I will discuss further later.

Financial support is another significant problem, because we have no permanent support. Some years, we receive funds from charities, but it is difficult to determine whether we will have sufficient funds to continue from one year to the next.

Finally, we have a problem with dissemination. We publish a report in a regular journal. I am not sure this the best dissemination method, and I will return to this issue later.

EXPERIENCE OUTSIDE U.S. II

D. Moher

DR. MOHER: To continue a point made by Iain Chalmers this morning, Dr. Freeson, the president of the Canadian Medical Research Council (MRC), asked me to provide a report on this meeting and how the MRC could participate in the trial registration process. The MRC is particularly interested in how partnerships can be developed within Canada, and other countries, to ensure the development of a registry process.

My comments today are based on my experiences as a long-standing member of a research ethics committee and as the Director of the Ottawa Stroke Trials Registry. My specific interests in the area of trial registration are in proposing a model, the Internet, which would allow access to trial registry information, and to suggest that ethics committees are the most important conduit to implement registration at inception of all trials.

The Ottawa Stroke Trials Registry (OSTR) was initiated in July, 1991. Its purpose is to collect information on all trials (in all languages) whose primary focus is stroke. This includes trials of stroke prevention, acute stroke treatment, and the long-term rehabilitation of stroke patients. The registry includes complete and published trials, complete and unpublished trials, ongoing trials, trials in the planning stage, and meta-analysis.

The OSTR is open to investigators involved in clinical stroke research, and investigators from around the world have registered trials with it. Currently, 52 planned and ongoing trials are registered. To register a planned or ongoing trial, investigators complete two forms. The first is a brief clinical inquiry form, which asks investigators to provide the name of the trial, an acronym, the name and address of a contact person, the primary objective of the trial, its recruitment status, and the name of the person completing the inquiry form.

The second form is a comprehensive four-page questionnaire, which includes a statement about the purpose of the registry. It then asks about the objectives of the trial, the eligibility criteria, treatment, and much of what Dr. Boissel presented. We also ask about the type of randomizations, type of masking, statistical analysis, type of validation information, source of sponsorship, the anticipated budget, and the name of the person completing the form.

We believe this information is important to patients, patient advocacy groups, granting agencies, investigators, and health policy groups. Recently, Bennett in the *New England Journal of Medicine* expressed concern about the under-representation of women in clinical trials. Similar concerns may be raised about other groups. Without a database of trials that includes this detailed information, it is difficult to evaluate whether clinical trials are fair, equitable, and inclusive of all groups in society.

Information that is added to the OSTR database is automatically updated every Monday morning at 4 a.m. The OSTR database is available on the Internet, an electronic superhighway that links individuals, organizations, and countries. This electronic linking is almost immediate, and free of charge. In Canada and the United States, the Internet currently is receiving a massive infusion of capital through the National Research and Education Network program and the Canary Project. This infusion of capital is very important, and could contribute to keeping the costs of implementing trial registration to a minimum. The Internet is free to us, and is likely to remain free for research and educational purposes for a long time.

The Internet has been described by many as the most significant development since the printing press. Over 30 million people use it in close to 100 countries, which may have important implications for access to information and technology transfer. On the Internet, one can access the complete Koran. In addition, the Guttenberg project is an attempt to archive over 10,000 public domain books in electronic text by the year 2000. The project has already published dozens of titles, such as the complete Sherlock Holmes mysteries, Alice in Wonderland, and the complete works of Shakespeare. All titles are freely accessible through the Internet.

The Internet is contains many resources for the health care community. There are books, newspapers, and commentaries on the Internet, but almost no information about current clinical trials. I am not aware of any comprehensive source which can be accessed for any information about planned and ongoing trials from around the world. This is a shameful situation that of which patients and the public are unaware but which would outrage them. We can and must do better.

The OSTR is available on the Internet. Those wishing to access information send electronic mail to an address indicating their desire to access the OSTR database. Once this request is received, two files are downloaded to the user explaining how to access and retrieve information along with a summary of file transfer protocol commands.

We will soon have a menu-driven system to access information on the OSTR database called GOFER that will enable users to avoid the technical details of accessing information. This is a free service developed to make the Internet more user-friendly. GOFER could be a very useful concept for the kind of registry we are discussing, as could the whole OSTR model. Internet could make registries accessible and available free of charge to interested parties from around the world.

Although the OSTR is a young registry, several individuals and groups have used it, including technology assessment groups, investigators, and systematic reviewers, and the National Library of Medicine. We have been asked to provide several sources of information to users, including the names and fax numbers of investigators conducting specific trials and information about the study design of all ongoing acute ischemic stroke trials.

I will use these five atrial fibrillation trials registered with the OSTR to provide an example of OSTR's usefulness. All of these trials addressed the question of whether aspirin or warfarin, compared to placebo, would reduce the incidence of stroke in patients with atrial fibrillation. They were conducted by five influential collaborative groups in Canada, Denmark, and the United States and were published between 1989 and 1991. At one point, all five trials were ongoing simultaneously. (Unfortunately, the Canadian Atrial Fibrillation Anticoagulation (CAFA) study was terminated early because of the results of the other four trials.) Combined, these five trials cost approximately $15 million and had a sample size in excess of 4,600 patients.

The availability of easily accessible trials registries would enable granting agencies to address whether five almost identical trials should be completed or whether investigators should be encouraged to collaborate on one large trial. This large trial might have a sample size adequate to address not only the primary hypothesis, but secondary hypotheses as well, thereby perhaps permitting monies to be redirected elsewhere.

Similarly, if trial registration were mandated and the resulting databases available on Internet, investigators would be able to access and transfer information about the technology from one country to another. Developing countries might be particularly interested in participating in trials in which the interventions are not costly and have important implications for their local health care needs.

In addressing other issues of technology transfer, 139 international stroke researchers were surveyed for their views of trials registries in general, and of the Ottawa Stroke Trials Registry, in particular. We achieved a 60 percent response rate. Ninety-four percent of the respondents supported the development of trials registries in principal. Eighty-five percent thought registries would foster and promote communication between investigators, and 84% thought they would provide an excellent means for keeping well-informed about planned and ongoing trials. I think the more user-friendly and accessible we can make registries, the more utilized they will become.

I would now like to discuss the role of ethics committees in trial registration. I believe they have a central role and are perhaps the most unique group to help implement trial registration. Much of the discussion in the literature involving publication has focused on the scientific issues. The substantive literature, which many of you have contributed to, has documented the existence of publication bias, factors that affect its occurrence, and the extent of the problem.

These issues are important, although I think they hide behind the much broader and more important issue of ethical irresponsibility. When patients participate in clinical trials, investigators have a responsibility to report the results. Without complete knowledge of all trials, we will be unable to address whether trials have been conducted fairly, equitably, and inclusive of all groups in society.

Unfortunately, some practices of ethics committees are strange and, perhaps, unethical. Their actions are often shrouded in secrecy. They go to great lengths to ensure that the protocol is kept confidential, often to the detriment of both patients and the scientific community.

What do ethics committees do to protect patients? In my opinion, very little. When a trial protocol comes before an ethics committee, members do not discuss whether there is evidence that the trial needs to be conducted, nor do they address whether there are assurances by the investigator that the report will be made public. I believe we have failed the patient in these areas.

We have also failed to carry out recommendations made by other groups. For example, the World Health Organization recommended that pharmaceutical companies make the results of their trials more available. However, in Canada, ethics committees make no such requests of pharmaceutical companies.

Despite this tarnished record, I believe ethics committees can play a central role in trial registration. They are in a unique position to promote, foster, and require trial registration, because all human clinical research must be reviewed by an ethics committee prior to initiation. Thus, these committees are the conduits through which to gather complete information for the registry at the inception of all clinical trials. This process applies in Canada, several European countries, and trials funded by the National Institutes of Health in approximately 450 affiliated institutions.

Ethics committees have an important role in educating investigators and patients about the need to register trials. They could help develop a patient information kit containing questions patients should ask before participating in a trial. For example, patients should be encouraged to ask whether the proposed trial has been registered and, if not, justification for this. They should also ask why the trial needs to be undertaken. They should request a brief, formal summary of the trial, as well. Investigators should be encouraged to document that a systematic review (possibly to include accessing clinical trials registries) has been conducted, the results of which indicate a need for the current trial.

In my opinion, to successfully develop a clinical trials registry, we need interaction between ethics committees and the Internet. The first step would be to consult with those involved in developing registry content standards, those familiar with the Internet, and those responsible for developing standards for ethics committees, (in Canada, the Canadian National Council on Bioethics in Human Research, and in the United States, the OPRR and the FDA, at a minimum).

This kind of interactiom would reduce or eliminate under-reporting of clinical trials, provide patients, patient groups, granting agencies, and health policy makers with important information to address, and make clinical trials registries up to date, user-friendly, and easy to access. This could be accomplished by requiring investigators to provide key information about their trials, in a standardized form containing relevant content for trial registries, along with their regular ethics applications.

In conclusion, we must move to implement trial registration. If necessary, this should be legislated. Registering all trials at inception can best be achieved at the ethics committee level. The Internet is an important model for helping to achieve trial registration. Several groups, including those involved with developing guidelines for registry content, those familiar with the Internet, and those involved with developing standards for ethics committees, should be consulted in developing trial registries.

Trials registries must be easy to access, up to date, and user-friendly. Funding should be provided to develop trial registration patient information kits, and also to develop electronic links between ethics committees and trials registries. Finally, the press has an important role in helping to educate the public regarding the under-reporting of trials, the importance of trial registries, and accessing information about trials.

EXPERIENCE OUTSIDE U.S. III

John Simes

DR. SIMES: I will present information on Austrialian Registries and report to the NHMRC, our Medical Research Council organization in Australia, recommendations of this meeting and ways we can collaborate in future activities and registries of trials.

I will begin by emphasizing the need for registries of clinical trial, including trials which are prospectively registered. We are discussing today the need for registries of trials, both the need for accessing all published trials, as well as obtaining information about studies which were registered in advance. In the past when examining trials without any particular therapeutic issue, one examined the published literature, identified all the studies from that source, and conducted a meta-analysis, or overview, of those studies. Usually, that exercise is extended to other, unpublished, studies to make the review as complete as possible. If, however, some of the unpublished studies were not identified, then the opportunity for publication bias exists. Another approach is to go to a trials registry, identify only those studies which were registered prior to their results being known, and undertake the meta-analysis or overview based on those studies.

Some 10 years ago, I examined trials evaluating treatments for advanced ovarian cancer. I identified 20 trials from the published literature that compared single-agent versus combination chemotherapy. A meta-analysis of those 20 studies demonstrated a significant survival advantage for the combination chemotherapy.

On the other hand, if one went to a registry of trials contained in CLINPROT (now maintained by PDQ), 16 trials were identified, of which half had been published. Meta-analysis of those trials failed to show significant survival difference. In this illustration, then, a different conclusion would be reached depending on whether the published or the registered studies were analyzed.

So in the future, meta-analysis is used to provide evidence-based guidelines for research, it would be preferable that, if we cannot ensure that we have identified all existing studies, we could analyze studies that are free of publication bias. To solve this problem, we need an international registry of all clinical trials registered in advance, with their objectives and outcomes clearly stated.

One of the problems in meta-analysis is study selection, i.e., selection of trials may have been influenced by the results. Another problem is that knowing the results of previous trials may influence the definition of the research question in terms of the intervention, patient groups, and outcomes. This may result in a different answer than a question defined without prior knowledge of some outcomes. In addition, there are problems with the quality of the data. Often the published literature does not allow

access to all randomized trials and all outcomes. Other problems exist with heterogeneity and the methods of analysis that it is possible to use.

Different investigators have reached different conclusions—albeit sometimes in terms of the magnitude of effect, rather than whether it is positive or negative—according to the studies they have elected to use and the subgroups they have elected to address. For example, a meta-analysis, or overview, of all cholesterol-lowering studies addressing primary or secondary prevention was undertaken by the Oxford trials group. In analyzing the odds ratios for coronary heart disease mortality, it showed that a small, but highly statistically significant, reduction in deaths from coronary heart disease with a slight excess, not statistically significant, for other causes of death. Overall, no significant difference in reduction of total mortality, albeit with a trend in favor of reduction in total mortality, was observed. Additionally, a meta-analysis published in the BMJ in 1992 by Ravnskov suggested a failure to demonstrate major benefits associated with cholesterol-lowering therapy; and cited other published meta-analyses as having been selective in including studies in that they excluded a number of studies he used for his endocrine therapies as a method to try reducing coronary heart disease. Other meta-analyses, at least one of which involved secondary prevention, have suggested there are significant reductions in total mortality associated with cholesterol-lowering therapy. Dr. Tom Chalmers was involved in one of those.

The conclusions of some meta-analyses may be controversial. This argues, when possible, for advance definition of the research questions with respect to such variables as the outcomes, total mortality or cause-specific mortality, the various subgroups, or, indeed, levels of baseline risk, as suggested by Davy Smith and Gordon Guyatt.

This illustrates a number of the ongoing studies examining cholesterol-lowering. Six of them, three in secondary and three in primary prevention, when combined will have increased statistical power in addressing issues of total mortality. Among them are some statin studies which produced larger differences in cholesterol than the earlier studies. Other studies now underway or about to begin will increase the patient sample size from approximately 30,000 patients to more than 60,000.

By combining these studies we hope, prospectively, to address some of the issues concerning cholesterol-lowering. In fact, two prospective meta-analyses are currently planned. One, called "3P," will examine the pravastatin studies in light of specific questions defined in advance. A second and a more ambitious undertaking, currently referred to as the Cholesterol-lowering Trialists Group, would involve more than 100,000 patients over the next 10 years (including the NIH Women's Health Initiative). Currently we have registered those trials and are in the process of defining the research questions. At present there is an agreement in principal among the various studies to continue to work together, but no formal agreement to undertake a prospective meta-analysis. That is our hope.

One strength of a prospective meta-analysis is that we can define prospectively the research question and the study selection criteria; i.e., the trials, patient groups, subgroups, and outcomes to be used. We will be able to obtain more reliable data—both more detailed data and agreed common datasets across the studies with standardization of outcomes. We will not resolve the question of heterogeneity but should be able to give a better explanation. Further, with the analyses being defined in advance, it can be more detailed by using time to event data an include an examination of predefined subgroups.

The analysis plan for the 3P study, and the model that we hope to use for the other one, is to find all the relevant trials *a priori* (to eliminate publication bias), undertake analyses by intention to treat (stratified or adjusted for trials, based on detailed time to event data), and define in advance which particular models we are using. We are considering a fixed effects model. A random effects model is only being considered for sensitivity analyses. Further, we will define advance subgroups, based on biological rationale and adequate statistical power.

I would like to address the issue of multiplicity of clinical trial data, and the problems this will pose, with access to the vast amount of information that will be available through registries of trials. We are familiar with the problems of multiple comparisons and the possibility that this can generate false positive results. There can be multiple treatments, outcomes, subgroups, repeated measures, multiple analyses of data over time, multiple methods of analysis, and multiple trials.

To take one example in terms of the subgroups, if we identify differences for one subgroup but not for another, we want to at least do that with the heterogeneity of treatment effect. We will find that subgroup more convincing if there is a good biological rationale, and if that subgroup was defined in advance, rather than a post hoc analysis of a database.

Advance registry of trials will not only enhance our meta-analyses but also help peer review of trials and help journal editors during review in determining, e.g. which particular subgroups were planned in advance and the *a priori* research questions. The registry of trials can also help avoid publication bias, with prospective meta-analysis, and in fostering collaboration.

I would like to comment on some activities of the NHMRC. One is the the quality of health care committee, which is involved in developing evidence-based clinical practice guidelines in Australia. The committee is interested in working with international groups to explore access to registries of trials is one approach for them.

The other example is the Australian National Cancer Network, currently being established to address improving cancer management and prevention into the 21st century. Some of the objectives are to identify effective interventions in prevention, early detection, and treatment, particularly from randomized controlled trials. For effective strategies, the registry will assess which interventions are cost effective and implement such strategies. Other objectives are to evaluate intervention programs by monitoring processes

and outcomes. For those whose effectiveness is not clear, we would undertake a preliminary analysis of cost-effectiveness and use that to identify future research to be undertaken, either by randomized studies or other means. We would like to take the approach of using randomized trials obtained through a registry of studies and a meta-analysis to identify effective or ineffective strategies.

One such registry of cancer trials is being established in Australia and is sponsored by the New South Wales Cancer Council. (Approximately 50 trial protocols have been submitted to us.) The purpose of this registry is to assist the New South Wales and other Australian cancer councils in funding trials research. It will help foster collaboration and avoid duplication of effort, keep investigators informed of ongoing trials, and (because a single country does not do sufficient trials for meta-analyses) forward trial details to relevant international registries to facilitate meta-analyses. We will provide an annual report on the type of cancer clinical trials research being undertaken in Australia.

The registry is collecting various details, two of which are worth bearing in mind in any register. One is the date the trial was registered, relative to the dates of recruitment. This is particularly valuable in cases where registries may contain studies not registered in advance, so that one can still select out prospectively registered studies.

Second, we are developing a critical appraisal worksheet for assessing the quality of studies, both to assist with the funding of studies of trials within the state for future activities, and also as a way of describing the type of work being done and providing that information to investigators.

I would like to close with a few recommendations that, perhaps, will receive less emphasis from other speakers. I support all the other things being said, but I thought I would just add these comments.

First, as we develop an international registry, it should include prospectively registered studies in advance of their conclusions. Second, that we consider that the purpose of the registries is not only as a basis for developing evidence-based guidelines, but also as a basis for planning further randomized trials. Finally, we should consider registering not only the trials, but also important questions; with the trials, interventions, outcomes, and subgroups that we consider of particular importance to be registered in advance, as well.

EXPERIENCE OUTSIDE U.S. IV

John R. Ennis

Department of Health, London

DR. ENNIS: I will describe our experience in developing a project registry system in England. Starting next year, this system will provide a consolidated database of health research and development, funded by the Department of Health in England and the regional health authorities.

The National Health Service in England was fundamentally reorganized in 1991 to separate the purchasing and providing functions which had been inextricably mingled. This led to a complex management structure which is now being reexamined, particularly at the higher levels affecting the regional health authorities.

Also in 1991, under Professor Michael Peckham, an NHS R&D strategy was estabshed for the first time. The three main objectives of that strategy are: an evaluative, knowledge-based culture; greater adoption of existing and new R&D findings; and better use of R&D resources. Project registries will probably make the most immediate contribution in this last area.

Within the R&D strategy, we analyzed information flows to understand the kind of information systems, including project registry systems, required. An information flow system summarizes our findings. We have users of information; projects feeding into project registries; the contents of project registries then being a resource for extracting from R&D, perhaps in the form of overviews; and back to the users again. There are other feed-ins, as well.

In my opinion, the uses of a project registry system can be divided into both retrospective and prospective categories. The retrospective use is accounting for publicly funded research. One prospective use is providing input to research overviews, particularly but not exclusively in the area of randomized controlled trials. Registries also support the efforts of those engaged in commissioning new research to target their efforts and resources effectively by avoiding unwanted duplication. Providing a framework for technology audits to those with responsibilities for technology transfer is another use. Enhancing patient accrual to individual trials is another objective. Finally, in the systems I will refer to, we considered but excluded the objective of using the project registry as a bibliographic record of published outputs of research.

Note that different categories of users have different interests in the system. This is important when considering a minimum dataset for a project registry system. Research funders probably have the widest set of interests in using a project register system, while health service managers are perhaps less interested. Researchers themselves may have a subset of the interests of funders and managers. Note that I have not included consumers as users of a project registry system. This may be an issue for debate.

The scope of the project registry system we have been developing was reviewed. Our focus is all "R&D projects which are taking place in or are otherwise of interest to the NHS." To be practical, however, we have initially focused upon projects being funded directly by the NHS authorities. "Are taking place," the present tense, conveys that this is a registry of ongoing research and that we have made provisions to include planned research, as well (although that raises issues of confidentiality). R & D projects, our scope, is wider than the clinical trial. While the exact meaning of "R&D" has been debated, I think it implies that the project must be generalizable in some sense. This is consistent with the view that the RCT is the best way to obtain evidence that is generalizable from the specific research context to a wider context.

The time horizon is "prospective from the first of January 1994," which is subject to confirmation. The geographic scope is initially, England, but we hope to develop links with the remaining three parts of the United Kingdom as soon as possible.

What information is collected? I think Dr. Harlan directed our attention to the concept of minimum datasets this morning. That is important because there is a trade-off between the volume of data one may attempt to collect and one's likely success in collecting it from the research community, not all of whom will see an immediate, direct interest in collaborating.

We have defined a mandatory dataset of five major areas: title, research objectives, methodology, lead researcher details, and project funding/management category. We are then assigning key words, using a thesaurus under development. Beyond the mandatory dataset is a dataset which is mandatory only for projects which the government itself is fully and directly funding. We have provided a further area of non-mandatory data items, for which a standardized format is being defined. Finally, there is a free text area.

Note that the 19 values we are using for methodology (and we have had the advantage of Carol Lefebvre's helpful input) allow the randomized control trial to be clearly distinguished from other types of methodology and design.

We are using MeSH as a basis for the thesaurus, but are adding additional terms reflecting the U.K. health care system. The "general practitioner fundholder," for example, is a term which does not exist in the U.S.. We also need to allow for the fact that our scope extends beyond the medical and includes areas of social and community care, as well. The British Library, the focal point for MeSH in the U.K., has been commissioned to develop this thesaurus. They will also be assigning key thesaurus words centrally, because this will be important in promoting consistent description of projects.

Finally, the regional health authorities are responsible for registering projects in their regions. With that in mind, we have suggested some sources of information that they might use in compiling their project registry. First are the continuing sources, which include major funders (such as government

agencies, charities, and industry); what I have called "invigilators" (local research ethics committees); existing subject-specific registries; and last, but not least, the informal human networks of research managers. Finally, there is one off-source which was a helpful baseline for us, and that is a 1991 snapshot of health services research in the U.K. done by Dr. Robin Dowie.

PATIENTS' VIEWS I

Cornelius Baker
Association of People with AIDS

MR. BAKER: First, I want to commend you for having included such a diverse patient perspective in these proceedings. We also strongly need the scientists to speak out loudly on the issues of clinical research for all communities of people who are affected by any disease to shares in our common knowledge and our common compassion for all humans. If scientists do not do that, then I think it will be difficult for activists to advocate effectively for your funding and research needs when speak to our governmental bodies in ways that scientists cannot.

We need to examine why the issue of clinical trials registries, at least from the AIDS perspective, is so important. With all due respect, Dr. Ennis, I think we have to point out that this is all about the patient. If we cannot get the patient involved effectively, and if the patient is not a co-facilitator of the on-going research, then in the end the research will be meaningless.

Particularly in AIDS communities, the issue of access to trials is a vital concern for women and people of color. Participating in a trial is dependent upon knowing about the trial, as well as on personal relationships, issues of geography, and finances. We need to recognize exactly who is being impacted by a particular disease and then explore the social and economic factors that affect how people learn about and access trials.

Note that this process so far has excluded more than women and people of color. The community of Atlanta, for example, has the eighth largest AIDS population in this country, yet has no clinical trials. To be effective, a registry must expand beyond geographical boundaries.

Another issue is understanding the trials and creating a broad-based language that we all can share in and contribute to. The idea of a thesaurus is an excellent one if it is written in a language that patients can understand and use to determine which trials are best for them. Allowing patients to be part of the screening process can help the scientists, as well.

What would patients want to know about a registry? First, they will want to know the intended objective of the research in clear and concise language. Also, while each patient participates with some expectation of a cure or an effective treatment, they will want to know the limitations of the research within the global context of their disease. Whether or not a particular clinical trial results in "the cure" for a given disease, participants should understand the contribution they are making to science.

It is also important for researchers to articulate what each patient is expected to do as a trial participant, along with the time committments expected. In addition, patients need to know what the

research team provides (e.g., travel reimbursement, other medical care). Patients should also receive an honest assessment of the intended target population and the researcher's experience in working with that population.

There are issues such as trial site locations, availability, and accessibility for patients. There are some more difficult issues. For example, if we create effective registries that are available to all communities and, in effect, equalize access to trials, can we also effect "blind" recruitment into trials and eliminate possible researcher bias in the selection of study participants?

Further, while a trial is ongoing, it is important that patients maintain constant dialogue with the research team so they are kept abreast of study developments and how they may fit within a broader scientific context. In addition to keeping patients informed, this will also aid in retaining patients in trials. It is also important to build patient advocacy into trials and to foster patient participation both within the ethics structure and with the research team itself.

Finally, we must not lose sight of who potential trial participants are. In the case of AIDS, for example, that means an increasingly poor community which may not have access to Internet. So, in establishing a registry, we must ensure that it is available to the majority of the population which is impacted by the disease, and which has traditionally been excluded from the research.

PATIENTS' VIEWS II

Patricia Barr

MS. BARR: I am an attorney from a small community in Vermont, the co-chair of an organization called the Breast Cancer Network (based in Vermont), a member of the Research Task Force, and the chair of the Grassroots Task Force of the National Breast Cancer Coalition.

The Grassroots Network of the National Breast Cancer Coalition has 250 member organizations and represents thousands of individuals across the country. We are committed to increasing research; improving access to high-quality screening, diagnosis, and treatment; and increased involvement of those living with the disease in legislation, regulatory process, and all aspects of clinical trial design.

The National Breast Cancer Coalition is also committed to an evidence-based health care system. We want a broader system of clinical trials that is accessible to all breast cancer patients. Our policy on breast health care includes concern that screening be done through a system of randomized national trials, and that we have a national system of trials. We believe we need to find the best screening time intervals for mammography for women between age 50 and 70. We also need to know the best screening system for women under age 50. Seventeen percent of breast cancer patients develop their disease before menopause. We also need data for women over age 70.

I am a breast cancer survivor. I was first diagnosed at the age of 37 while still nursing my nine-month-old daughter. Although the surgeons I consulted assured me that this was an early diagnosis, during surgery it was discovered that I had more than one tumor and 14 of my 16 taken nodes were positive. Thus, it was not unexpected that I recurred two years later. I have been in remission since that time, for the last four years.

As a member of an organization offering peer support to other women with the disease, and as a coalition board member with direct responsibility for working with a nationwide grassroots network, I continue to have contact with cancer patients at time of their diagnosis, at time of treatment choice, and during their treatment. In addition, I serve on the regional facility's institutional review board (IRB). As you have heard, this is comparable to the ethics committees in Europe. We are charged with reviewing all protocols offered as a part of randomized clinical trials by the hospital medical staff, before any volunteers are enrolled.

My regional hospital is unusual, because although located in a rural community, we are affiliated with the Dana Farber Cancer Institute through the Clinical Community Oncology Program. This NCI program enables patients at small hospitals to participate in clinical trials. As a member of the IRB, I regularly review protocols offered through two national research groups, funded by the NCI.

I have raised the issue of services that are not available in our hospital and the mechanism for referral, and I asked whether we had an obligation and ethical responsibility to refer patients, regardless of disease, to sites where clinical trials would be available. However, there was no mechanism within the hospital committee structure to consider such an issue.

When I was first diagnosed six years ago, participation in clinical oncology trials was not an option at that medical center. Unlike most rural women, I had the resources and energy to travel to Boston for my surgery and subsequent treatment. At a Boston teaching hospital, I had access to protocols in trial, and was able to participate in two research programs—one identifying a new tumor marker and the other studying relaxation and stress reduction during treatment. It is a tragedy and a lost opportunity that only two percent of breast cancer patients in this country are enrolled in trials. I understand that over 30 percent of breast cancer clinical trial participants worldwide are from Japan.

We know that progress in our search for cures and improved quality of life will come only through the slow process of research and clinical trials. Part of the explanation for the small percentage of trial participants is the number of physicians and patients who are without easy access to information about trials in progress. Even if their facilities do not participate in the Clinical Community Oncology Program, knowledge of trials might lead to referrals to institutions where trials are available.

As a nation, with the medical, scientific and patient advocate community taking the lead, we must change the attitudes and misunderstandings that surround human volunteer research. Extending the reach of available information is an important component of attitude change. While we look for cures, we must also look for life-extending and quality-of-life improving protocols. Increased participation in clinical trials will lead to increased knowledge. Increased participation is the only way we can hurry time at this point. We will not be able to shorten the appropriate length of any particular trial, but we can extend the number of trials that are ongoing if we develop the means to increase registration.

A registry available to doctors, patients, educational groups, and advocates will improve the quality of care available to all patients, regardless of their own participation in those trials. The registry will eventually result in increased clinical trial participation. Without knowing what trials are being done, doctors cannot recommend participation to patients and patients cannot learn about available choices, or even express interest in participating in trials.

As an example, members of my small support group in rural Vermont range in age from 30 to 70 and come from across the economic spectrum, rich and poor. Many of our members have raised the issue of weight gain during breast cancer treatment as women on Megase and Tamoxifen gain a significant amount of weight. The oncology office we are associated with, although a member of the

community oncology system, has had little success in gathering information about the cause of the weight gain or what can be done to control it. We would like to evaluate existing information in a systematic way and then work with the oncology office to design the best possible support program. However, finding appropriate published research and identifying relevant clinical trials is not now possible, even using MEDLINE. Clearly we need a registry to help doctors, and women like those in my group, find the information they need. Currently, we must rely on drug companies alone for information. I would much prefer to consult a clinical trial directory that is indexed for professional and lay access.

I am aware of individual research facilities attempting to expand their participation base through computer networks and bulletin boards. These efforts tend to be regional. To provide the best patient care, we must have a national system. We must not be satisfied with existing registries or methods of distribution of the current registry information. We know many doctors will not use PDQ online. We need a system that provides updated hard copy and a map of all CCOP participating facilities. Research and clinical trial programs are arduous to design, time consuming to implement, and often require a long wait for useable results. How can we make reasonable funding decisions unless we have a well-designed national and international system for tracking all ongoing research, including clinical trials? We all know that resources are limited and that groups representing one disease do not want to be fighting with those representing other diseases. No matter how successful advocacy groups and researchers are in increasing available resources, choices between projects to be funded will have to be made. The registry is an essential tool for making good judgements.

Patient advocate organizations can be strong allies to the scientific community. We want evidence-based health care and will continue to push for more trials, easier access to trials, and the broadest possible registry of research. Many advocates are trial graduates. As patients, we have a vested interest in a strong research system. We know the benefits of trial participation, the feeling that one is making a contribution to medical advancement, the sense that one is getting special attention. We can be the clinicians' and researchers' messengers in the community.

PATIENTS' VIEWS III

Tricia McCauley

MS. MCCAULEY: I am the coordinator of the National Alliance for Oral Health, a consortium of non-profit and professional organizations sharing an interest in oral health needs as they apply to special patient populations. Our membership represents individuals affected by TMJ disorders, cleft lip and palate, Paget's disease, Sjogren's syndrome, ectodermal dysplasias, and other genetic and systemic disorders. Our organization endorses establishment of clinical trials registries. Many constituents are affected by disorders about which little is known. Clinical trials often become their only chance for establishing a nearly normal lifestyle. Even the smallest advances can lead to major life-changing opportunities. Many of our member organizations actively support clinical trials through funding and/or access to patients. As patient support groups, we have a responsibility to know about clinical trials pertinent to our constituents. It is equally important for us to be assured that all clinical trials we endorse are safe and efficacious. Our organizations are in an excellent position to recruit patients for clinical trials they endorse. Newsletters and other publications effectively bring information on such trials to those who need them most. Best of all, we enable our constituents to do something very special—to help themselves. Those involved with clinical trials are often the beneficiaries of extraordinarily good care, and others are saved from making the mistakes of what is found to be ineffective.

Clinical trials take place in a number of venues. Certainly we are all familiar with the Clinical Center at the NIH, but extramural programs and private industry trials may be less familiar. At times, duplication of effort may take place because one group is unaware of a similar project being done elsewhere. Because a centralized registry does not exist, information on successes and failures may not be readily available to scientists or patients. Clearly, such registries would be advantageous to science, investigators, clinicians, and patients.

Establishing clinical trials registries could eliminate another problem that has become clear to our member organizations. Recently a clinical study was undertaken by an investigator at the NIH Clinical Center. A variety of patients were seen, each undergoing about a week's worth of tests. Invaluable data was collected on these patients, some of whom were affected by extremely rare disorders. Unfortunately, the investigator moved to a different institute, and the work now lies dormant under the jurisdiction of the sponsoring institute. Had a registry been in place, perhaps another investigator might have taken an interest in the project and continued it.

The success of clinical trials is our only real hope for the future, and the National Alliance for Oral Health hereby underscores its belief in the importance of clinical trials registries to benefit scientific knowledge, those dedicated to its quest, and those patients desperately awaiting the results.

DISCUSSION

DR. GREENHOUSE: I am with George Washington University's Biostatistical Coordinating Center and formerly an Associate Director of the National Institute of Child Health.

Dr. LaRosa, I recently read that the National Organization of Women and other organizations objected to the 1993 FDA guidelines on the participation of women, saying that they were too restrictive. They argue for the participation of pregnant women in trials with drugs, which is the most outrageous thing I have ever heard.

DR. LAROSA: What I say will be very cautious. A number of groups strongly advocating that. In fact, the publication of our guidelines in the *Federal Register* was supposed to occur tomorrow, December 7, but has been postponed for a day or two while Congress made sure the policy stated that we do not exclude women of child-bearing potential in these trials or studies. We are talking about clinical research, including clinical trials, so that means the complete spectrum.

I think this must be handled very carefully. In the next several months we expect to release an Institute of Medicine report on the ethical and legal issues of including (or excluding) women in clinical studies. That will further explain our thinking in this area.

I share your concern about this. However, if we do not put women of child-bearing potential, note I do not say pregnant women...

DR. GREENHOUSE: I think they should be separate subjects. I am talking about pregnant women, not those of child-bearing years.

DR. LAROSA: I understand. I think that will not be done with these guidelines for a long time, and there are some extraordinarily knotty issues that must be worked through.

DR. TEMPLE: I want to comment on the participation of pregnant women in control trials. The routine exclusion and failure ever to include pregnant women results in some bad outcomes. A strong case can be made that where drugs are likely to be used in pregnant women, for example drugs that treat high blood pressure, it is worth studying them. We recently had an experience in which ACE inhibitors, a benign class of drugs for most purposes, proved quite toxic to the fetus in the third trimester. This was a big surprise, because that is not where one ordinarily looks. There seems little doubt to me that it would have been preferable to have had trials that explored the possible consequences before the use of these drugs became common practice.

Also, we do not do a good job with registries of exposure of pregnant women. So for someone to say we ought to do it better is quite a good point.

DR. SILBERMAN: I am from the GAO. We spent much of the day listening to why we should have a registry, and I will restrict my comments to a registry of ongoing trials. I think the real issue

is, if you are going to have a registry, what problems does it pose, and how do you best deal with them? I do not think these issues have really been discussed today.

I think the major reasons for a registry are:

- to alert patients to treatment options. That is a valid and reasonable objective, but we must deal with the crisis of raised expectations on the part of patients, and the idea of randomization and not getting the treatments they want.
- to increase recruitment. We must deal with the fact that we have had PDQ now for a number of years and it has not increased recruitment.
- to help overviews. I think this is really one of the strongest arguments for the registry because it will reduce publication bias. But it may, in fact, lead to a different form of bias, which I will call implementation or execution bias—the likelihood that the study comes to fruition, and is done with studies where the envelope gets opened early and/or where there is crossover that is unannounced.
- to increase oversight. How, as a research community, do we deal with a Congress that is becoming increasingly active in directing research, both in subject and in terms of who becomes recruited?
- to what extent does the registry become the rationale for not funding any research in an area where we have already had one study, in this time of tight budgets?

DR. SIMES: I disagree with the idea that another trial is not needed in situations where several trials have already been done. This also relates to a point made by Tom Chalmers earlier about whether the other studies are necessary once we have clear evidence. I think that, in cases where studies have clearly demonstrated a potential for a therapy to be effective but where there is still uncertainty, further large randomized studies are both ethical and reasonable. An example is the ISIS-2 study referred to earlier.

And the other comment on that, is that not all randomized trials undertaken could demonstrate that a meta-analysis was an effective therapy and hopefully leading to that being more widely used in practice always follows, is the recent ISIS-4 trial. This analyzed on of three treatments including magnesium. Meta-analysis in that study had shown fairly convincingly (in terms of a p-value) that the intravenous magnesium for acute infarction was of benefit. However, the ISIS-4 trial reported last month showed no benefit of intravenous magnesium treatment. Thus, I think undertaking a large study to confirm very promising results of a meta-analysis may be reasonable, rather than concluding that we have a definitive answer.

DR. SILBERMAN: I think my comments may have been misinterpreted. I am an advocate of the registry, as I think most of the people in this room are. However, we must be alert to the problem

that funders may decide not to fund further studies on Subject A because we already have a study in that area and may, instead, decide to fund studies in Subject B.

DR. GREENHOUSE: Some of these comments will seem to be hostile toward registries, but they are not. Paul Meyer and Jerry Caulfield and I go back some 25 years, to the inception of clinical trials here at the NIH, in thinking that we need sources where other clinical trial data have been entered. Our thinking at the time was primarily for scientific purposes. What I have heard today is an expansion of those purposes in many directions, such as using registries to help journal editors and for administrative purposes. However, every time one adds a purpose for a registry, one also adds to the content and the cost will increase.

For example, the Heart Institute might have 30 different disease areas in its program. Would we have 30 different registries? If each one required perhaps one-half million dollars for development, that would eliminate support of clinical trials for the next two years in the Heart Institute. This would hold true for other institutes.

Next, I am concerned about the expression of deep concern for patients, as if that concern has not existed and is not already a part of current clinical trials. Data safety monitoring committees are required by both NIH and the FDA for every major Phase III clinical trial. This committee's primary objective is to moinitor adverse effects for patients. Trials have been stopped when death and other serious adverse effects developed.

Who is going to take the stand for clinicians and others who do clinical trials? Let me give you an illustration. I am involved currently in a trial in benign prostatic hypertrophy. Many people know of all major ongoing Phase III clinical trials. They do not need a registry to know what is being done. One clinician, however, expressed the point that he is seriously worried about spending the next four or five years on the trial where his authorship will be lost in a group authorship in the *New England Journal of Medicine*, and another investigator will use the registry, do a meta-analysis, and receive all the honor for discovering that benign prostatic hypertrophy is, or is not, in some sense improved by the agents that involved. We must be concerned about that.

DR. SILBERMAN: I have a question that not have an immediate answer to but ought to be considered: As the effort to automate the patient record through health care reform proceeds over the next 3 to 10 years, what integration, if any, might there be between that and an automated trial registry?

MR. BAKER: I share the concern on the issue of duplication of research, because duplication can be desirable if there is a clear justification of the objectives and a clear outline of what populations were included in the original research. The Concord example is one where it is good that we did not

force earlier collaboration and coordination, because that may have resulted in a premature end to both trials.

I also think that the language of scientists has led us to believe (in many cases because of ego, papers, et cetera) that the patient's interest has not always been first. Hopefully, conferences such as this helps us begin to work more closely together.

Finally, some registries are now being devised, in fact the District of Columbia is about to establish one that is a patient registry, where the names entered would be accessible to the clinical trial workers. We are leaning toward opposing such registries. We think that they could easily be abused, could result in patient bias in recruitment, and could result in a total loss of confidentiality, particularly for patients with HIV as opposed to AIDS, who could experience a greater abuse of their rights.

DR. LIONEL EDWARDS: I represent the Special Populations Committee of the PMA. Judy LaRosa, when you said that substantive differences would require subpopulation analysis of different minority groups, it depends on which minority groups, and what do you define as a substantive difference? It depends very much on the disease. Could you comment?

DR. LAROSA: As I said, I prefer that you read the policy thoroughly when is issued. Much information is buried in the policy from NIH, and we tried to be sensitive to the scientific community. I am not necessarily referring to clinical trials. Rather, somewhere within the process of clinical research, one must consider the question to be answered, the existing data, and the principal investigator, to make the case for substantial data.

Now, if you think I am waffling, I am. We at the National Institutes of Health have a huge task to help the scientific community, the IRBs, and others to interpret these guidelines. That is why we have allowed at least a year for comment, because we know that it will be a shift in the way we do business, and we want to ensure that we are being as sensitive and thoughtful to the scientific community as we can. This is a message that we hope everybody here will take back to the scientific community.

The guidelines will apply differently in different cases. One example that is used frequently involves low birth-weight babies. This varies tremendously in the Hispanic population alone, I am told, so one must examine the data across the subpopulations. However, in other instances, that may not be necessary. If you can provide substantial evidence that there are no differences, the study section will, presumably, understand that.

DR. TEMPLE: For what it is worth, awareness of the full range of studies in a particular area may help you to explain why any given study need not answer a question entirely.

DR. LAROSA: Exactly. That relates to another point we tried to address in the guidelines—the scientific portfolio of a particular institute. For one reason or another, it may be impossible to have all

groups at all times in all studies. For some compelling reason, someone might tell us they are going to do a particular trial with particular groups. The institute might then choose to do another trial with other groups in another location.

What we are trying to say is, come to us with the best reasoning, and work with us as we put this in place and try to resolve the problems. Congress is absolutely firm on this point. I will tell you that we must plea-bargain down from valid analysis in all clinical research.

DR. TEMPLE: There will still be a problem, and it will not be resolved by anything you do, when you get to subsets of minorities.

DR. LAROSA: I appreciate the nuances.

DR. TEMPLE: You may, but you are talking major efforts here. On an optimistic note, since 1988 we have asked companies to take their complete new drug application database and analyze it for differences among demographic and other subsets. Crude as these analyses are (because some are not even based on randomized trials), they do show certain kinds of differences that one knows exist. For example, in any database from an ACE inhibitor, it is easy to discern a difference between black and white responses. This is predictable. We have also had recent examples in which side-effect profiles in women and men were different. Again, not in a surprising way, but because women were smaller and got a bigger dose. It is not that we discovered anything major, but these relatively crude analyses can detect big differences if they exist. I do not know if that will help solve your problem with Congress, but it is some reason to hope that if one knows the whole database, one can learn something from it.

DR. LAROSA: Precisely. That is why I made such a point about gaining information as we go along. I have been at the National Heart, Lung, and Blood Institute when two national education programs were started—the high blood pressure program and the National Cholesterol Education Program—and I know that it takes a long time to get a shift in scientists. Also, as the advocates carefully point out, one must be able to recruit and retain people in these trials and studies. It is very important that we realize that this will require a period of adjustment.

It is also important that we obtain these data to eliminate asking these questions when in a Phase III clinical trial in 5 or 10 years. We will know whese things are necessary. We need to be patient and open-minded as we proceed. We are in this situation because we have not collected these data over time, and I believe the Congress is correct on that point.

DR. IRWIN MARIN: I am with Emblematics. Another message I think the scientific community needs is a word of caution about meta-analysis. Many statisticians would not agree that meta-analysis is as wonderful as presented at this meeting. We should remember that there are other techniques. There is a danger that, if we codify a methodology and collection of data in one particular

way—for example, saying thresholds were fixed when they may be changing over time—we might encounter difficulty both in including new data included and in analyzing it in a different way to get the wrong conclusions.

EXAMPLES OF U.S. REGISTRIES: PDQ

Ann Thurn

DR. THURN: I work at the International Cancer Information Center, where the National Cancer Institute's information services, including CANCERLIT, a bibliographic database of publications on the topic of cancer, the *Journal of the National Cancer Institute*, and the cancer information database known as Physician Data Query (PDQ), are maintained. I help ensure that the content of the PDQ database is kept up to date.

The International Cancer Research Data Bank was established by an Act of Congress in 1971 to promote the rapid and effective exchange of cancer information throughout the world. Those at the NCI responsible for organizing, updating, and disseminating this rapidly growing collection of knowledge decided to establish a computer-based information system that would complement NCI's other mechanisms of scientific information dissemination.

The first attempt became available in 1974 as a database known as CLINPROT. CLINPROT was a "computerized database [that summarized] the investigational treatment protocols sponsored by the NCI." It was formatted to allow distribution and searching on the National Library of Medicine computer system. Because knowledge of the National Library of Medicine's specialized searching language was required, a search intermediary such as a librarian was often needed for an effective search.

The protocols in the database were "not indexed by stage, histologic type, participating investigators, or institutions." As a result, CLINPROT was not well suited for clinical care decision-making. Instead, it served as an online catalogue of clinical research, providing physicians who were designing clinical trials with information on trials performed by others. It proved to be of generally limited use to health care providers.

In 1982 CLINPROT was replaced with the PDQ database. PDQ is also available through the National Library of Medicine but is a completely menu-driven system, so no specialized searching skills are required.

The PDQ main menu lists the range of information included in PDQ. The three major components are: cancer information statements on treatment, supportive care, and screening. Directories which contain listings of physicians who treat cancer patients and organizations that offer comprehensive cancer care are available. The protocol section contains active and closed clinical trials, including NCI sponsored and voluntarily submitted protocols.

All trials in the database are cancer-related. Most analyze cancer treatment, but trials on supportive care, screening, and prevention are also included. Almost all trials involve some form of

intervention; however, some purely observational trials have been included. There was a Stage I trial, with Stage I testicular cancer for observation after surgery, and there are trials under supportive care that examine the psychosocial effects of having cancer.

Not included in the database are, for the most part, epidemiologic type studies. For example, the early studies on the effects of smoking, which were purely observational, would not be included here. Also not included are in vitro studies, such as the ancillary laboratory studies that go along with our clinical trials, and those that examine prognostic indicators or pharmacokinetics alone.

The database currently contains information on approximately 1,500 active clinical trials, about 10 percent of which are from foreign countries. It also contains 6,000 closed protocols. Because the results of many of these trials are not published, this part of the database provides a useful resource for details of unpublished clinical trials.

The database can be searched for either active or closed protocols alone, or for both. Standard therapy protocols for the treatment of a variety of cancers are also included.

NCI-sponsored trials are entered as soon as they are approved. As mentioned, PDQ also includes trials submitted voluntarily by researchers. The database is updated monthly and, on average, 35 percent of the active protocols are revised each month.

When a voluntary protocol is received, it is checked for completeness. For example, does it contain an entry criteria in the statistical section? a treatment plan? an informed consent form? If the document is complete, it is reviewed before being entered into the database by members of the editorial boards that have been organized to ensure accuracy.

The criteria for entry of a voluntary protocol into PDQ are fairly lenient. As long as the protocol design is based on rational scientific information, is not unduly harmful to patients, and will answer a scientific question, it is included. Approximately 20 new voluntary trials are added to PDQ each month.

The protocol document itself is used as the source for the information included in PDQ. The document is summarized, the summary is approved by the protocol chairman, it is indexed, and put online. About 60 new protocols per month are added. Another 50 are closed each month and moved to the closed protocol section of the database.

The monthly updating and revision of 35 percent of the active protocols is very labor intensive. To ensure the information in the database is up to date, a systematic review process is followed. Once the initial protocol summary is reviewed for accuracy and approved by the protocol chairman, the summary is re-reviewed by the protocol chairman or designee on a regular basis. Non-NCI-sponsored clinical trial protocol summaries are reviewed three times a year, while NCI-sponsored trial summaries

are reviewed once each year. This is effective because we receive all the amendments for the NCI-sponsored trials as soon as they are approved by the NCI Cancer Therapy and Evaluation Program.

The list of participants for each trial is updated six times a year for most trials. The status of each trial is updated monthly. One large cooperative group, the Eastern Cooperative Oncology Group (ECOG), provides this update information electronically. We receive hard copy information for all other clinical trials.

The protocols are indexed by a number of terms which may be used individually or in combination to narrow a protocol search. For example, to retrieve the protocols in PDQ for melanoma, one might first find how many protocols are indexed by this diagnosis.
That would result in a list, and options would be provided for protocol retrieval by diagnosis or cancer type. The search results show that, currently, 95 active clinical trials are listed in PDQ for melanoma. Depending on how one plans to use the results of the search, one can opt either to view these results as-is, or to further narrow the scope by any of the criteria listed in the menu. I will briefly discuss some examples to show how the interactive process functions. One can narrow retrieval by modality. Any of these modalities can be selected—including biological response modifier therapy, chemotherapy, phototherapy—either individually or in combination. The phase of the clinical trial and the geographic region are practical ways of narrowing a protocol retrieval for a specific patient.

As mentioned, the database can be searched for either active or closed protocols. The default for protocol status is active, but it might be useful at times to include closed protocols.

Once satisfied with the retrieval, the next step is to review the results. The protocol display options menu provides a choice of seven predefined display formats. These range from BROWSE (which lists the titles and provides the option of selecting any items for further display), to LONG display (which provides all the information contained in the database on that particular trial). The CUSTom display option allows users to design their own display output. Many data elements are available for each trial listed in PDQ. Users may select to review any of these available data elements.

The database was designed to include a broad range of information, because the usefulness of different types of information could not be predicted. The designers wanted to ensure that physicians could find the trials for which specific patients with specific disease profiles would be eligible, and that patients could determine the options available for their particular situations. When available, citations and abstracts for clinical trials are included and can be added to the display at the user's discretion.

We hope, by this summer, to allow users to narrow their searches by eligibility criteria, so that searches can be tailored to the needs of individual patients. For breast cancer, for example, the user

would be given the option of using the following criteria and limiting the results of the search: axillary lymph node s[illegible] menopausal status, estrogen and progesterone receptor status, and performance status.

Besides being accessible through the National Library of Medicine, PDQ is available through timeshare vendors. This mechanism has increased PDQ availability worldwide; there are now more than 10 vendors in the U.S. and Eur[illegible]e. CD-ROM has made the local copy available at a fixed cost for large volume users. We currently have about 600 subscriptions. The Cancer Information Service makes PDQ information available free of charge to anyone who calls their 1-800-4-CANCER number. Parts of PDQ are currently available through Internet and fax. However, we feel at the moment that the interactive search process by either these mechanisms is unweildy. We would, however, be interested in doing more of this in the future.

Our usage information indicates that PDQ use has been steadily increasing. Because of the variety of mechanisms for disturbution of PDQ, compiling this information is difficult. We know that the Cancer Information Service, our biggest single user, receives about 30,000 requests for information on clinical trials each year. For 1992, this represented approximately 23 percent of all the calls to the Cancer Information Service.

We are currently working with the Division of Cancer Prevention and Control to answer the following questions: Who uses PDQ and through what mechanisms? What sections do they use? How frequently do they use it? What impact does it have on cancer care? We hope this will help focus our efforts to provide useful information to researchers and health care providers.

EXAMPLES OF U.S. REGISTRIES: AIDSTRIAL/AIDSDRUGS I

Debbie Katz

MS. KATZ: I am from the National Institute of Allergy and Infectious Diseases (NIAID). I will discuss the AIDS Clinical Trial Information Service (ACTIS).

In 1986, NIAID initiated its clinical trials program for AIDS. From the start, the Institute received criticism from various AIDS constituency groups and the general public. They were concerned about the difficulty of learning where NIAID and NIH-sponsored clinical trials were conducted and who eligible participates were. In the Fall of 1988, NIAID conducted a user survey to determine remedies for a major communication problem. The overwhelming recommendation from a broad range of individuals was that a centralized, very current, telephone-based system was needed. Shortly thereafter, Congress passed specific legislation in this regard as part of the Health Omnibus Program Extension Act, or the HOPE Act of 1988.

To quickly and efficiently establish a telephone network, NIAID entered into an interagency agreement with the CDC to expand the CDC's established AIDS Clearinghouse to include a clinical trials information service. Through this mechanism, the program began as the NIAID AIDS Clinical Trials Information Service (NACTIS) on May 9, 1989, within six months of the first discussions with CDC.

This mechanism also allowed for major cost savings. The concept of a coordinated effort, as well as cost savings, was the force behind adding the FDA database. For the FDA, it was a requirement of the HOPE legislation. Eventually, the National Library of Medicine was included. By changing the Service from the NIAID Clinical Trials Information Service to the AIDS Clinical Trials Information Service, the Public Health Service was (and still is) able to provide consumers with one-stop shopping for information on most clinical trials for HIV. We increased access to the database by making the data available to the National Library of Medicine.

Before November 1988 when the HOPE Act was enacted, the FDA was legally prohibited from releasing any information on drug or biologic products to the public. Passage of the HOPE legislation lifted this legal restriction from HIV-related therapies for which clinical trials testing efficacy were underway. For the FDA, the HOPE legislation described the type of trial to be included in the database and specified what information could be released, including information on trials to test efficacy, protocols released with sponsor permission, and treatment INDs.

The FDA then developed and published in the *Federal Register* a policy statement for use in defining efficacy trials under the HOPE legislation. According to this statement, the FDA will consider a protocol to be a trial to test efficacy if it is not on clinical hold and meets the following

requirements: the trial is at a certain stage of development, the design is such that the trial is considered adequate and well controlled, the trial examines sufficient endpoints, the trial is of sufficient size to support a claim of efficacy, and the trial was designed to yield information that could be used for drug approval.

In addition to trials that are required to be released, the FDA requests that sponsors help carry out the intent of the HOPE legislation by permitting it to include information about any AIDS-related product under development. Thus, between the information provided by the FDA and the information provided by the NIH, ACTIS lists a majority of all clinical trials for HIV infection.

The AIDS Clinical Trials Information Service can be accessed by calling a toll-free hotline number (1-800-TRIALS-A), TTY/TTD, international telephone line, and fax. The AIDS Clearinghouse can be called with a question about trials and be switched over to ACTIS. There is also online access through NAC.

Calls are answered by health care professionals. They are able to provide both verbal and written information responses to inquiries, in English and Spanish.

EXAMPLES OF U.S. REGISTRIES: AIDSTRIAL/AIDSDRUGS II

Gail Dutcher

MS. DUTCHER: I am with the National Library of Medicine. In July 1989, the AIDSTRIALS and AIDSDRUGS databases became available publicly on the NLM MEDLAR system. This made them accessible to thousands of online users, both directly and through intermediaries such as medical librarians. Health care providers and others from around the world can search these databases. If they wish, they can access them through GRATEFULMED, the library's software package that enables untrained users to easily search our databases by filling in a blank on their computer screen.

The AIDSTRIALS database, which has the protocols, currently has about 540 trials, some 130 of which are open. Each record in the AIDSTRIALS database is a description of a single clinical trial, while each record in the AIDDRUGS database contains detailed information about a single agent being tested.

Starting with AIDSTRIALS, I will briefly review some of the major types of data included in each database—those considered core to the project.

Each trial has a descriptive title that provides some basic information. There is also a field with the purpose or description of the trial. For NIH trials (where we get the entire protocol for summarization), there are paragraphs describing the purpose, rationale, and basic methodology. Spanish language versions of this summary information are created for the ACTG and CPCRA protocols that are multicentered and have large Spanish-speaking populations within their regions. We have not yet added these online at NLM, but they are available through the toll-free number.

The information in the description and purpose is written at a level appropriate for educated consumers and health professionals. This is the primary piece of the record that is sent out by ACTIS when patients, family, or friends request a printout.

Some of the most important information contained in each record concerns eligibility. We have split eligibility into two major categories, inclusion and exclusion criteria. Each category contains specific information that allows one to identify whether or not a patient qualifies for a specific trial. Among the elements of inclusion criteria are specific requirements for the state of the disease, including clinical manifestations and ranges of values for a number of laboratory tests. Sex and age requirements are also listed. Details about current and prior treatments, both required and permitted, are also included.

To some extent, the exclusion criteria are the mirror image of the inclusion criteria. They provide specific values for what is not allowed under the study and what makes an individual ineligible. These include age and sex, as well as prior and current treatments.

The trial status, indicating whether the trial is open or still accepting patients, is very important. This field has three levels: open, closed, and completed. This is one of the most significant fields for searching if the purpose is to identify a trial to enter or to refer a patient. We have no plans to purge trials from the database, even after results are published. The database will retain value as an historical record since July 1989. In addition, since not all trials result in publications, this may be one of the few records of those trials.

One of the most-searched fields in the database is the trial locations. This provides essential information about where the trial is being conducted. Users regularly ask first for a trial in their own locale. In addition to the institution name and address, contact information—generally a local telephone number—is also provided. We also try to keep current with the accrual status of individual locations and whether each site is still accepting patients. It is possible that the trial is listed as generally open for accrual, yet individual trial sites may have stopped adding new patients.

Most of the trials included in the database are drug trials. Each trial record contains the generic name of the one or more agents being studied. We use the generic name because each drug may be referred to by many different names and we had to choose just one. Any other names that we know about and can identify are listed in the AIDSDRUGS database. Also included in AIDSTRIALS is the unique number assigned to each drug in the AIDSDRUGS database, to make it easier to move to that database and locate the record that fully describes the drug.

We also add two types of indexing to the records. We have about a dozen categories, and there are some trials that do not fit into any of them. The categories were originally assigned to make it easier to find specific groups of trials that might otherwise be difficult to retrieve. Over the 5 years we have worked with these databases, MeSH has added new terms, making some of our categories obsolete. Because people have become accustomed to searching with them, howewever, we are not going to remove them.

After the trial records are extracted and entered into the database, a trained indexer uses NLM's medical subject headings (MeSH) to assign index terms. Although there are many access points in the record, this additional indexing makes it easier for information professionals who are often the intermediaries in online searching, and supports some of the advanced features of our online retrieval system. It also facilitates access through GRATEFULMED, NLM's microcomputer-based software program for easy searching.

There are numerous other fields, as well. These include both expected and current accrual, the trial sponsor, and a myriad of dates, such as when the trial was entered into the database and when different fields were updated.

The AIDSDRUGS database is an essential companion to AIDSTRIALS. Rather than repeat the information in the record for each trial in which a drug is used, the information is all placed in one record in the drugs database. Each record in this database describes one single drug or agent. Naturally, since many of these are experimental agents, the amount of information available about different agents varies widely.

We include all of the generic, trade, and any other names and synonyms that have been used for each agent, including the chemical and systematic name. This list of names provides a vital access point, since a single drug may be referred to in multiple ways. We also include the registry number assigned by the Chemical Abstracts Service. This is useful for pointing to information in other sources. However, this number is often not available right away for new agents and biologics.

We also include pharmacology, drug interactions, and major adverse effects, if known. We have also found it useful to provide a simple phrase describing the type of agent the drug is classified as, such as an antifungal agent. Users are then able to locate all the agents in that category and can search the trials database for all trials using that type of agent.

Finally, we include a selected list of 10 or 12 citations about the agent from the AIDSLINE bibliographic database. We update this list every 6 months. Someone wanting a comprehensive bibliography about the agent, or a highly specific search on a single aspect, would have to do his or her own search. We also have additional fields of a variety of chemical and physical data. The information for AIDSDRUGS comes from published, peer-reviewed sources, including journal literature, handbooks, and other compendia of drugs and chemicals. Information is also extracted from the protocols, when no other fully reviewed and published data is available.

I will turn now to how data is added to the AIDSDRUGS and AIDSTRIALS databases. As Ms. Katz stated, the two sources of our data are NIAID and FDA. Once NIAID obtains FDA approval for a study, they provide the full protocol to ACTIS, where the relevant information is extracted, abstracted, and entered into the database. FDA, in order to protect the proprietary information included in protocols submitted to it, prepares its own synopsis and provides the summary to ACTIS. This is done through the Center for Biologics and the Center for Drug Evaluation.

This information is then entered into the appropriate fields in the database, and the database record is indexed. Each time a new trial record is created, the AIDSDRUGS database is also updated. If a new agent is being tested, an entire new record is created for AIDSDRUGS. However, even if one record for the agent already exists, it is at least updated with the new protocol number and any other new information.

Although FDA does not actively participate in the AIDSDRUGS database, records for any new agents that occur in the non-NIH trials are created and added to the database. Since we do not have

access to the protocols, only public sources are used to compile those records. Once the records are created, they are immediately available to the reference specialists answering the telephones. Data is sent to NLM for updating the online database every 2 weeks. We also can and do update more frequently, when very significant trials are announced.

It is important not only to add new trials as they begin, but also to keep existing information up to date. The most changeable item is the status of the trial. All NIAID trials are tracked by a central contractor who verifies trial status at least mo ly. That central coordinating site provides regular updates to ACTIS, where it is incorporated in AIDSTRIALS. Updates to privately sponsored trials are solicited at least monthly by ACTIS staff, through letters and follow-up telephone calls. We try hard to ensure the accuracy and validity of the information we provide, but it is very difficult to keep up. In addition accrual status, other parts of the records may change. For example, a part may be added to or deleted from the protocol, a drug dosage or a regimen may be modified, or location information may change.

It is also important to maintain the AIDSDRUG records. As trials progress and are completed, more information becomes available on drug interactions, pharmacology, and adverse effects of drugs being tested. This must be incorporated into the drug records.

On a simpler level, even new names must be added as they occur. Constant vigilance is required.

The initial focus of this project was on treatment and intervention trials, since one of the main purposes—particularly for NIAID—was to improve trial accrual. Early on, Phase I trials were not entered into the database. That changed quickly for the NIH-sponsored studies. Recently, epidemiology and natural history studies were added. These clearly are not randomized clinical trials. However, since AIDSTRIALS and AIDSDRUGS serve several purposes, adding these trials is useful. First, it is very difficult to obtain this information because there is no central resource, as we have been discussing. By adding this to AIDSTRIALS, we are developing a single source for information about AIDS-related clinical studies.

Also, health care providers have indicated there are patients who may not be prepared to participate in or eligible for a clinical trial who may be interested in some sort of involvement. This listing of epidemiology and natural history studies gives health care workers the opportunity to refer their patients or clients and helps diversify the participants in these studies.

FDA remains bound by the legislation and, with few exceptions, only efficacy studies from the non-NIH sponsors are entered into the database. As said, Phase I and epidemiology studies were excluded initially, but are now included. For the most part, FDA is still adhering to its efficacy algorithm, thus non-efficacy studies from private sponsors are excluded, though not prohibited. We

like to have those in the database and, when one of the sponsors offers us the information about the trial, with FDA concurrence, we add them. We have not been recruiting behavioral, social science, or psychosocial research. However, if sponsors of such studies recommend their inclusion and provide us the data, we will consider adding them.

Finally, I will address usage. As you know, we provide multiple access points. I will focus on the telephone service and the online system. The use of these two channels varies greatly.

At the toll-free telephone number, bi-lingual reference specialists respond to calls from around the world, but primarily from the U.S. The telephone service has been averaging 3,000 calls a month within the last 6 months, or over 35,000 calls a year. Of these, about 75 percent were from patients, their families, and friends, and another 10 percent came from the general public. Approximately 13 percent were from health care providers.

The electronic situation is quite different. The average number of online codes accessing the AIDSTRIALS database each month is between 120 and 150. These are different codes. It is not possible to say how many users there actually are, since many of these are institutional codes, where the number of users can vary. We have been averaging between 15 and 20 hours of online searching each month, which is a very small number compared to a bibliographic database like MEDLINE. The type of use is very different, with users doing very simple searches on a drug, a location, or a few items within the eligibility criteria fields. There is rarely a need for recurrent searching, and no need for monthly updating when new information comes in, as is the case with bibliographic databases. There is usually one trial printed out and a patient referred. As far as we can tell from examining the user codes, the majority of users are either health professionals or health-related institutions, such as medical libraries, pharmaceutical companies, and hospitals.

In September 1993, we made the entire AIDSDRUGS database and records of open trials from AIDSTRIALS available as text files on our FTP server and through the Internet. Those connected to the Internet can quickly and easily transfer these data files to their own computers. Since the retrieval software is not part of this, however, they must provide their own means for searching or using the data. This was done in response to requests from people running AIDS bulletin board systems and providing information to community-based organizations. We know that a number of these groups have been downloading the data and then sharing it with many other bulletin board systems and users. In addition, ACTIS provides data to many community-based organizations for their own printed directories.

Finally, we have been working with the Canadian HIV Trials Network (AmFAR), the Pan American Health Organization, and many other groups to develop an expanded international registry of HIV clinical trials. Trying to get these international groups to work well together has proven

difficult. We are, however, very satisfied with the cooperation we have achieved in developing AIDSTRIALS and AIDSDRUGS. Each of the four Public Health Service agencies involved provides a different perspective and expertise, and we are all very proud of what we have been doing.

EXAMPLES OF U.S. REGISTRIES: OMAR REGISTRY

John Ferguson

DR. FERGUSON: The OMAR (Office of Medical Applications of Research) database began in 1974 and was discontinued in 1979 because of a lack of funds. It was started again in 1985 and included more observational studies. In 1988, it was restricted to controlled trials, though institutes could add uncontrolled studies. Its purpose was to assist Congress, the Department of Commerce (because of the Stevenson-Wydler Technology Transfer Act), and other government agencies. Information from each institute was collected annually and managed by OMAR.

In 1989, there were 1,400 active trials. (For various reasons, NCI trials were excluded from the OMAR database, but information on them is available in PDQ.) Information is available as a printed report and online at the NIH on the WILBER system, making access difficult.

The definitions have evolved since 1978. Randomized controlled trials were added with the possibility of using historical controls as part of the definition, in 1988, 1989, and 1990.

The datasets included have also envolved, with different items having been included and excluded, arbitrarily. This is one reason we would like to modernize this database and, perhaps, move it elsewhere, depending on the outcome of this workshop.

The current dataset includes the administering organization (usually the institute), the principal investigator, the project officer, the Information for Management, Planning, Analysis, and Coordination (IMPAC) project number, title, starting date, last date of projected support, study status (planning, open, or closed), age range of population, projected sex/size of the sample, type of intervention, type of control, amount of support, and amount and type of support (which is often difficult to ascertain). We are reorganizing this.

EXAMPLES OF U.S. REGISTRIES: VA REGISTRY

Yick-Kwong Chan

DR. CHAN: I am from the Department of Veterans' Affairs (VA) and will give an overview of the VA's registry. Before I describe the registry, I would like to give you some information about the Veterans Administration Cooperative Studies Program. The program was organized in its current form in 1972 and is centrally directed by the Chief of the Cooperative Studies Program from Boston and Washington, D.C. Four coordinating centers provide the study design, data management, statistical analysis, and administrative support. They are located in Hines, Illinois; Palo Alto, California; Perry Point, Maryland; and West Haven, Connecticut. A fifth center, the clinical research pharmacy, is in Albuquerque, New Mexico. It is responsible for drug-related activities such as developing drug-handling protocols, negotiating with pharmaceutical firms, and packaging, distributing, and accounting for drugs.

The Cooperative Studies Registry evolved from the Master List, which was created in 1972 for administrative purposes. It included items such as study number, status, date, name of chairpersons with phone number and address, and the coordinating centers. In the early 1980s, the coordinating centers were required to submit semi-annual reports describing the major features of each study under their management.

In 1990, Dr. David Wise of the Perry Point Coordinating Center created a database containing 95 studies in 23 medical areas. It contained basic descriptive information on each studies. The current registry was initiated in December 1992 to provide a more detailed, standardized description of VA cooperative studies.

The registry comprises all clinical trials funded under the VA Cooperative Studies Program, including studies in active planning. Exceptions include studies by individual researchers participating in multicenter groups not under the auspices of the Cooperative Studies Program, and those conducted by VA Cancer Study Groups sponsored by the NCI.

The subject areas included in the registry are multidisciplinary. Most studies are conducted in the U.S. within the VA, and joint ventures with the NIH may include Canada. The registry encompasses randomized clinical trials and observational studies from 1972 through June 1993. As of 1993, there were 140 studies in the registry with 15 in planning, 43 active, and 82 completed. There are 27 medical categories for studies in various areas such as cancer, cardiovascular, hepatic, infectious, and psychiatric.

The core content of the registry is almost as described in "The Directory of Registries of Clinical Trials," by Dr. Easterbrook in Statistics in Medicine, 1992. Additional features such as study

outcomes, major manuscripts, cost effectiveness, and potential benefits and impact were included. The registry printouts contain a table of contents followed by the individual study description. There are two indexes: one by disease category and the other by study status and study number. The registry is updated centrally and continuously with the new biannual reports, reprints and publications, and revised versions of the study reports from the centers. The entire registry or a portion thereof may be requested at no cost, as a printout or on diskette, in writing, by fax, or by telephone.

CORE CONTENT OF REGISTRIES I

Kay Dickersin

DR. DICKERSIN: At the University of Maryland, we are compiling a "registry of registries." If we have not already contacted any of you who are associated with registries, please let us know. We would like to add you to our list and contact you on a regular basis in order to disseminate information. Kristen Larson, who is here today, is working on that. Also at the University of Maryland, in conjunction with the Cochrane Collaboration (we are the Baltimore Cochrane Center), we are establishing an international registry of published trials. Eventually, this will be linked with the NLM.

I will now briefly present the core content of trials registries, mostly from the perspective of the international group. The concept of "core content" means that all trials registries collect the same kinds of data. Each registry can collect data specific to itself, but all registries should collect some similar and overlapping data.

There are four reasons to establish a core content. First, the usefulness of registries is related to their content. Second, why people want to use registries should determine their content. For example, if we are interested in drugs that are being tested, we should collect information on interventions being tested. Third, many small registries with the same core content could be merged into one large registry covering many different subject areas. Finally, collaboration toward common goals has many benefits.

In 1991, we formed the International Collaborative Group on Clinical Trials Registries. All registry keepers are invited to join this group, along with anyone else who is interested in the topic. We have held annual meetings since 1991, in conjunction with the Society for Clinical Trials.

Let me outline the objectives of this international group. The first is to inform one another of our current activities and to compare experiences. One topic of disucssion, for example, is how we keep our registries up to date and how they are accessed. We want to establish an international network so we can keep apprised of what others are doing and can improve our own registries.

Another objective is to prepare and maintain an up-to-date directory of registries, which is what we keep at the University of Maryland. We also have been considering a core content of registries, proposing actions to stimulate new registries, and helping those who want to start registries. Finally, we are seeking ways to obtain funds for starting new registries.

The recommended core content that the international group are all trying to keep includes trial name, a registration number (a numbering system was suggested to us by Maria Lebron of the Online Journal), principal investigator's name, disease or condition, objectives, method of patient selection,

test treatments, comparison treatments, design (parallel or crossover study), method of randomization (for example, envelope or telephone call), outcomes (main, secondary, and additional), whether central laboratories or reading centers are used, dosage regimen, duration of treatment and follow-up, number of centers, sample size, start date, operational details, termination date, the various centers (coordinating center, reading centers, etc.), funding source, publications, and dates of enrollment for first and last patients.

Basically, this core content is very similar to what has been described already by a number of the registries presented. What I think is particularly exciting is that there is an international group collaborating on registration and, once we decide to establish an international registration system, we should be able to move forward quite quickly.

CORE CONTENT OF REGISTRIES II

Jean-Pierre Boissel
University of Lyon

DR. BOISSEL: Once the content is defined, the next question is how to fill in the various boxes. There really is a series of problems there, and we can call upon the ethical boards, investigators, coordinating centers, sponsors, and drug regulation agencies for help.

I would like to show part of the results of the FICHTRE project, which was funded by the European Union to explore various issues in clinical trial registration. I will discuss the part of this project focusing on the institutional review board (IRB) as a source of data for a planned and ongoing clinical trials registry.

We sent more than 300 questionnaires to IRBs in Belgium, Greece, Denmark, France, Germany, Italy, United Kingdom, and Sweden. When I left Lyon, 70 questionnaires had been returned. The key question on the questionnaire was, "Would you agree to communicate data to an outside registry?" Seventy percent of those responding said they would agree. Perhaps most interesting were the reasons given by the 30 percent who refused. Most refused because the law in their respective countries makes it impossible for them to disclose confidential information to an outside party. This is especially true in Belgium and France, where ethical committees are bound by special legislation. A few said they would refuse because they believed it is not the responsibility of the ethical committee to participate in a registry of planned or ongoing clinical trials, despite their understanding that it would help them in mediating protocols.

The majority of those who agreed to communicate to an outside body wanted the registry to be supported by public funds and preferred that the registry be maintained by a university. However, the majority also wanted to limit the content of transferred data to the name of the sponsor, the name and address of the Principal Investigator, and perhaps the type of drug tested and control treatments.

In conclusion, provided that European directive allows or mandates the single committee, the IRB can be the most convenient source for planned and ongoing clinical trials registry data. For those not familiar with how the European Union functions, a "European directive" supercedes the laws within European countries. Thus, if a decision is made in Brussels that the ethical committees in the twelve European countries should disclose data to an outside clinical trials registry, they are obliged to do that, regardless of their own country's laws.

DR. FERGUSON: I am not sure we could do that within the NIH, much less the United States.

CORE CONTENT OF REGISTRIES III

Lois Anne Colaianni, M.L.S.
National Library of Medicine

MS. COLAIANNI: I am with the National Library of Medicine and want to make some comments about the core content. Some of these comments have been made already, but I want to reemphasize them. First, we need to carefully think through the purposes the registry will serve, because the registry's purpose will govern the data it should contain.

Second, each data element must be clearly and unambiguously defined, so that the data attached to each field is understandable by all. This is especially critical in a case such as this, where cooperation between various agencies and among different countries is involved.

Third, one field should provide the bibliographic citations to connect each clinical trial with the literature published about the trial. We must be able to move among databases with some reliability, and should seek to avoid the problems that have been encountered with searching MEDLINE.

Fourth, the data elements in existing registries should be used to the extent possible. This will not always be possible, however, and I am concerned that people and institutions may not be willing to change their ways of doing things to accommodate the needs of the registry.

Fifth, the data entered into each field must be readily available and easily obtained. The more difficult it is to obtain a piece of data, the less likely it is to be reliable. It is also important that there be an authoritative source for all of the data, and that the data be kept up to date.

Sixth, the registry should be linked to MEDLINE. This will permit people to search the literature, identify clinical trials in the registry, and obtain the registry number to easily locate additional citations to publications about a particular trial.

One should not underestimate the effort involved in developing the registry's core content. This is not an exhaustive list. Additional ideas for items to be included have been presented at this meeting. I recommend that we approach the task by breaking it down into parts, such as administrative data needed for maintenance. For example, it was mentioned that over 10 percent of all of the MEDLINE records are changed each year. As that database grows, you can imagine the amount of work involved in updating it. Other categories might include data elements that would be required for patient referral, data about substances being studied, data for linking with MEDLINE, data for prospective referrals, and others as appropriate.

DR. BALAS: In proposing a standardized set of items for the registries, are you considering the numeric results or just other characteristics of the trial? The numeric results are critical for any kind of synthesis. However, they are often not available in publications. However, requiring that type of information means one cannot register many trials without contacting the authors.

DR. FERGUSON: Do you think it would be advisable to include the numeric results?

DR. BALAS: I would say yes, but it would probably limit the number of people who would be interested in the registries because it would require much more effort.

DR. SIMES: Including the numeric results of the trials has advantages and disadvantages. One of the disadvantages is that investigators are reluctant to include results before they are published. Once a study has been published, then, people will be more inclined to include the results. An advantages would be that the results would be hlepful in doing meta-analyses.

Even if the results from unpublished trials are not included, listing those studies in the registry can be helpful in interpreting a meta-analysis of the studies for which the results are available. One could say, "We have the results of eight of the 10 trials we know about, and the conclusions of a meta-analysis of those studies is as follows." Then, perhaps one could speculate about the results of the remaining studies. Without the use of those registered trials, however, one would lack knowledge of additional studies that might change the conclusions. Thus, registering studies without results can still be very helpful.

DR. FERGUSON: I agree with that. Also, an issue that arose in our clinical alert workshop 3 years ago concerned the need for an editorial or peer review process that is informative and at least slightly digestive when making results known. I think that was the problem that NLM was wrestling with on their clinical alerts.

DR. SUSSEL: I am from the Canadian HIV Trials Network, and there is an issue that we grapple with all the time that has not been mentioned today. Perhaps it is not an issue in the context of the NIH, but we gather clinical trials information for AIDS and our biggest hurdle is that the pharmaceutical industry will not give us trial information for reasons of confidentiality. I do not know if anybody else has had that experience, but it is a problem.

DR. FERGUSON: We will hear from the pharmaceutical industry tomorrow morning, and perhaps we can enter that discussion then.

DR. TEMPLE: My question is for Tom Chalmers, and Iain Chalmers if he is still here, because they work from published literature all the time. It has been mentioned that access to information about unpublished trials might allow a more informed meta-analysis of existing trials. My question is,

how important is that factor? I would guess that it is mostly the smaller trials with little difference between treatments that do not get published. They are not likely to have a major effect on the meta-analysis. How important a component of the reason for having a registry is the ability to do better trials?

DR. T. CHALMERS: One may worry about how unpublished trials might alter our conclusions, but we have yet to find a situation in which they really have. I am looking at John Simes now, and also at Iain Chalmers, and although they may bring the p-value from .04 to .06, unless one worships the p-value, that is meaningless. We have not found a situation in which the unpublished trials have changed the clinically applicable conclusion that either there is a tendency for something to work, or it is clear that it does not work. We have done this by comparing meta-analyses done by different people, in which others have used the unpublished data and we have not. I think this is because the unpublished trials are usually small and when added, do not seem to change the clinical meaning of the meta-analysis.

The main point is that, in every example I have seen, there has been a difference in the statistical significance when unpublished data were used, but there is no statistically significant difference between the two, suggesting that they really do come from different populations.

DR. FERGUSON: Meaning that the trial was done and was not published, and the ones that were done and published come from different populations? Is that what you mean to say?

DR. T. CHALMERS: No evidence that they do come from different populations.

DR. FERGUSON: There is no evidence that they do come from different populations.

DR. T. CHALMERS: There is evidence that they may switch the p-value from .04 to .06.

DR. I. CHALMERS: I have reached a slightly more cautious conclusion, which is that we are still in the early days of doing systematic reviews of randomized trials, and I think it is difficult to come to a confident conclusion about the importance of those that are not published. One possibility for the discrepancy between the overview of magnesium trials in myocardial infarction published by Tio and others, and the results of the ISIS-4 trials, is that a systematic search for even published trials was not thorough enough to unearth some of the small trials of magnesium, let alone the unpublished ones.

The other example that concerns me is giving aspirin to pregnant women who are at increased risk of developing hypertension or of having a growth-retarded baby. The early overviews are encouraging, but we know that there are trials that have not been published. The biggest trial that has been published, an Italian one, has not confirmed the early results of the meta-analysis. We know also that a trial has recruited 10,000 women without stopping on the basis of any direction from the data monitoring board. So I think it is too early to make any confident general rules.

DR. FERGUSON: John Simes, if I understood correctly what you published several years ago, it said that including the unpublished work did make a difference.

DR. SIMES: Overall, I share the same caution that Iain Chalmers expressed. While one would get the same answer most of the time, having large numbers in the individual trials will result in a slight overestimation of the size of any effect, particularly if the analysis includes several small studies.

DR. T. CHALMERS: It is too late in the day to get into the details of the unreliability of the very large trials. In answer to Iain Chalmers, however, it is clear that the early trials of magnesium—which showed a remarkable reduction in mortality—could never be obviated by unpublished trials. It would require hundreds of trials sitting in file drawers with an opposite or negative effect to make up the eight or nine that were published. In that situation, the difference is in the administration of magnesium and is not a difference in the publication of large versus small. In the interest of time, however, let me agree with Iain that we do not have all the data we need.

DR. FERGUSON: So you are not recommending that those unpublished trials stay in a file drawer and be lost forever?

DR. T. CHALMERS: No, I think we need all the information we can get on the subject. I have not, however, been convinced that it is worth devoting a tremendous amount of effort to the unpublished trials.

DR. MARIN: I am from Emblematics. Perhaps what is called for, from a scientific point of view, is a theoretical study that looks at the robustness of the meta-analysis or some sort of perturbation on the structures, to see in general when something stays within the confidence bounds of the structure, so that one could accept it.

DR. MOHER: I think we must make it clear that, when patients participate in trials, the results should not remain in someone's file drawer. At the very least, we must come out and say that. We have an obligation to make what is done available. Once it is available, one can decide whether it is good or bad quality, to be included or not.

DR. TEMPLE: Perhaps some trials are not published because their importance is not perceived. That is, the meta-analysis looks at a question that was not what they were trying to find out originally, and what they found out originally did not strike them as worth bothering with.

The idea that a meta-analysis that is positive at .05 is okay, and one that is not quite at .06 makes me nervous. I like Richard Pitot's approach, which said, "If it is not at .001, do not bother with it." There is so much uncertainty in meta-analyses and the quality of the trials that are in them, one ought to set a higher standard for believing them in the first place. I continue to wonder whether

a few small unreported trials are likely to turn a meta-analysis that is persuasive at that level. Perhaps not everybody agrees that this cautious early attitude is warranted, however.

DR. GRAY: I would like to raise an issue about the content of registries and propose that the principle be "simplify, simplify, simplify." One needs to strike a balance between completeness and complexity.

DR. T. CHALMERS: I would like to quickly second that with a little historical data. Between 15 and 20 years ago, we tried desperately to set up a registry of clinical trials but failed, because we tried to include so much information that it became totally impractical to record it. Had we begun with a very simple registry, we might have been way ahead of everybody.

REGISTRY MAINTENANCE AND QUALITY CONTROL I

Lois Anne Colaianni, M.L.S.
National Library of Medicine

MS. COLAIANNI: I would like to discuss five issues concerning registry maintenance. First, there needs to be an authoritative source (usually the funding agency) to collect and review the quality of the data before it becomes available in print, online, or however it will be disseminated. That entity also needs to review the data on the system in case changes occur after it has been input. Database creators, such as the National Library of Medicine, usually do not have the expertise to provide this kind of authentication, because their staffs have not been trained for this task and are not as familiar with the data as the funding agency or the principal investigator.

Second, it is important that the data be available in a timely manner, depending upon the use of the database. A core number of elements may be used in common by registries. Then, as different purposes are served, additional data elements would be added. Databases that provide patient referral, for example, could be input or updated more frequently than those that simply registry that a trial is underway or planned.

Third, despite best efforts, there will always be errors in the database to be corrected. Proper provisions must be made to carry out this maintenance activity.

Fourth, the thesaurus must be kept current. For example, we add 600 to 700 different medical subject headings (MeSH) to our thesaurus each year. The subject access to the registry must be updated as MeSH changes occur to facilitate proper searching.

Finally, let me reemphasize the need for data elements that are agreed upon and commonly used. This will allow people incorporate data from a registry or other source and into another registry. Closed trials reported in the literature would have to be re-annotated with bibliographic citations.

None of these issues is insurmountable, but they do require a commitment and resources. If a registry is to be a dependable resource, both money and staffing must be made available for proper maintenance.

REGISTRY MAINTENANCE AND QUALITY CONTROL II

John C. James

Division of Research Grants, NIH

DR. JAMES: I ran an inventory for 5 years, between 1975 and 1979. At the time, the inventory was published in two volumes. The Division of Research Grants accomplished the computerization of the registry with the help of a contractor. I would like to discuss some of the issues and problems we encountered in automating the system.

We used a one-page questionnaire containing 23 data elements. These included a standard abbreviation of the ICD, the grant contract or intramural project number (which allowed us some degree of quality control over administrative information on the performer), Prinicipal Investigator, dollars awarded, NIH project officer, clinical trial title, trial initiation date (month and year), expected trial duration (in years), amount of NIH support by fiscal year (including current and projected support, and cumulative support to date), study purpose (a concise statement of the hypothesis being tested—this was the most essential scientific description of each of the trials), trial population age range, and number of participants.

The form was printed on legal-sized paper in fine print. The instructions appeared on the reverse side. We refined both the inventory and the instructions as time went by.

By the fifth year we were running smoothly and paying our contractor $25,000 or $35,000 a year. We then encountered three significant problems which lead to the demise of the registry. First, our division's workload increased while our staffing did not. Second, we encountered resistance to the inventory on the part of some of the institutes. Third, we were forced to open the contract up to competitive bidding. Other problems we encountered included survey forms being returned months beyond the deadline dates, and printing problems when producing our annual narrative summary of each year's collection effort.

All of this leads me to ask whether it might be feasible to address the inventory problem by adding new tags in the IMPAC system. This is our fiscal and administrative computer system, which mostly tracks the process of peer review of grants and grants management. Much of it is dollar oriented, while there is some science in such things as the titles of grants and the study section that reviewed them.

We have a separate information system, called CRISP, which is a scientific indexing system with a controlled thesaurus of about 10,000 terms. Several years ago, upon request, they added a new term called clinical trials or clinical studies. Unfortunately, it became diluted to the point where it no longer met a very tight definition of a clinical trial.

Recently CRISP modified its thesaurus and redefined a clinical trial as "a controlled study designed to assess the safety and efficacy of new drugs, devices, treatments or preventive measures in humans by comparing two or more interventions or regimens." I believe the research documentation section, which runs CRISP, would be willing to further refine their definition to comply with the wishes of a central committee or group. It then might be possible to add marks to characterize grants deemed to be clinical trials with an indexing term called "clinical trial," permitting them to be retrieved from the CRISP system.

Other special issues including DNA research, AIDS, and nutrition have had special tags put in our information systems. With AIDS, for example, we created a new data element in IMPACT which the institutes can either accept or override at the time an award is made. This can create problems when the grants management specialists are not familiar enough with the science of the projects to determine which should and should not receive the markers.

If we were to have clinical trials markers in CRISP and IMPACT, another problem would involve the complexity of the mechanisms for supporting clinical trial activity. These can range from one grant/one trial, to very large grants with multiple trials, to one trial with multiple contracts, to cooperative groups with multiple protocols and participating grantees. Thus, a very complex type of reporting must be dealt with.

At a minimum in developing a computer-assisted activity of this type, we would need the ICDs to enumerate and name each clinical trial, provide the funding amount, and identify the supporting grants and contracts as a checkpoint for each clinical trial that meets the definition of the central committee.

REGISTRY MAINTENANCE AND QUALITY CONTROL III

Joan Porter

Office of Protection from Research Risks, NIH

DR. PORTER: I will discuss institutional review board (IRB) review and approval criteria, and suggest how these relate to a clinical trials registry or registries. I want to begin by talking about the Office for Protection from Research Risks, and the IRB system as it operates in this country. The Office for Protection from Research Risks (OPRR) is the component of the Department of Health and Human Services that acts on behalf of the Secretary in implementing the Department's regulations for protecting human subjects.

In the United States, many institutional review boards operate under the authority of the Federal Policy for the Protection of Human Subjects. Also called the Common Rule, it consists of 15 sets of identical regulations for the protection of human subjects. These are based on Department of Health and Human Services regulations which went into effect in the 1970s and have been revised several times, most recently in 1991.

The Office for Protection from Research Risks has negotiated what are called "assurances of compliance" with the regulations with about 450 major United States institutions. Most of these institutions have agreed to conduct all of the research under their auspices in accordance with the regulations of the Department of Health and Human Services, regardless of the source of funding. This extends the impact of the regulations considerably. While there are about 450 institutions, there are over 500 IRBs that operate under those assurances. (Some institutions have more than one IRB.)

The Office for Protection from Research Risks has negotiated many other assurances for single projects, and for multiple projects under a limited program, all of which require institutional review board review. Other Federal departments and agencies negotiate assurances and require institutional review board review for research involving human subjects that they support or conduct, as well. In addition, the Food and Drug Administration requires institutional review board review for research involving drugs or biologics or devices which it regulates. The IRB basic review and approval criteria are the same for all Federal departments and agencies conducting or supporting research in the United States under the Federal policy, and for the Food and Drug Administration. Thus, there is a widespread network of institutional review boards in this country using basically the same criteria for review and approval.

In approving research, an institutional review board asks the proposing investigator to provide information indicating that risks to subjects are being minimized by the use of procedures consistent with sound research design, and which do not uneccessarily expose the subjects to risk. Registries that

present data about risks of drugs or procedures are, therefore, quite helpful to institutional review boards. The AIDS drug registry discussed yesterday is an example of an online and up-to-date information source in which risks can be identified.

Further, institutional review boards are interested in equipoise. That is, if there is a doubt about the effectiveness of an intervention, are there sufficient data to show that there is no question about whether or not the intervention works, or whether one intervention is more effective than another? An IRB would, theoretically, question why it is necessary to expose individuals to risks or discomforts of additional research, or to less effective interventions, if the research question is already known. Inasmuch as data concerning completed trials are available and used by investigators in framing research questions and making the case for the need to conduct further study, the institutional review board can be reassured that the research question still needs to be addressed.

The extent to which institutional review boards insist on exhaustive searches of literature or registries probably is quite variable. Similarly, the extent to which IRBs have faith in meta-analyses probably varies, as well.

In theory then, institutional review boards could tell a proposing researcher that the question has already been answered; that there is sufficient data to warrant no futher investigation. Indeed, both Drs. Chalmers presented compelling examples in which conducting meta-analyses and having access to additional information might have resulted in some trials never having been conducted.

It is always difficult to say that something is redundant. When can one generalize from one study to another? Is there a different population? Is there a slightly different question? The existence of a registry and the analysis discussed yesterday might be quite helpful. The IRB has a mandate to ensure that risks to subjects are reasonable in relationship to anticipated benefits, if any, and that the importance of the knowledge that may be reasonably expected to result is clear. The extent to which investigators can clearly and comprehensively cite current data about risks and benefits enhances the ability of IRBs to protect subjects. In offering these data, clinical trials registries can play an important role in this process.

Still another mandate of the institutional review board is to ensure equitable selection of subjects. That means that the institutional review board is concerned with making sure that some groups are not inappropriately burdened with risks of research by their inclusion. It also means that groups or persons are not inappropriately excluded from trials either, resulting in injustice. Direct or indirect exclusion of groups results in failure to identify differential responses, and may result in inappropriate generalization of research results to those populations.

Clinical trials registries that provide data indicating there are reasons to believe, or not to believe, there are no substantial differences in gender or ethnic and minority populations are useful in

considering recruitment criteria and sample size. Clinical trials registries that indicate that sufficient numbers of racial or other ethnic minorities and women can be included at other sites where a research trial is being conducted may also be helpful to an institutional review board considering approval of research. Dr. LaRosa has already suggested that this might be an important consideration in knowing precisely what numbers and types of people are being enrolled at different sites.

Further, if registries of ongoing trials are available to consumer groups and physicians, there is some reassurance to institutional review boards that recruitment goals are enhanced, that equity of access is heightened, and that there can be responsiveness to potential participants ineligible for a protocol who might be eligible to enter other research trials with different criteria. Some of our presenters yesterday from patient groups emphasized that quite clearly.

The NIH guidelines soon to be issued in response to the NIH Revitalization Act, about which Dr. LaRosa spoke yesterday, includes a statement that institutional review boards will also adhere to the policy. Exactly how they intend to ensure this is an area that is bubbling with anticipation and speculation. Institutional review boards will need data, directly and indirectly available, about the inclusion of minorities and women in trials.

Yet another criteria for institutional review board approval is that the research plan should provide adequately for monitoring the data collected to ensure the safety of subjects. It can enhance the research plan for safety monitoring if an investigator or data safety monitoring board can access data relating to the risks of the research interventions from similar ongoing trials or published trials.

The institutional review board is also required make certain that the informed consent information and process are appropriate. For example, in addition to a discussion of risks, benefits, and alternatives in the informed consent, the IRB may require a statement that significant new findings developed during the course of the research which may relate to the subject's willingness to continue participation be provided to the subject. Again, depending on what data are ultimately contained in the clinical trials registries, this element of protection of the subject might be enhanced by reference to registry data provided in an updated consent process.

To summarize, the use of clinical trials registries, both of completed and ongoing trials, by investigators could be of considerable importance in presenting data to institutional review boards to help them make decisions about protection of subjects.

The next question is, should IRBs require the use of clinical trials registries by investigators? In large part that depends on whether or not the institutional review boards are aware of the registries, and whether or not they have confidence in the registries as a source of data. Another factor is the extent to which the data can be reasonably obtained elsewhere.

What role should insititutional review boards take in becoming the primary source for entering data into the registries? I would stop far short of the suggestion, made by two presenters yesterday, that IRBs should be the primary source of data entry. I say this for several reasons.

First, the regulations under which institutional review boards operate in this country were created in the 1970s, when the research enterprise was considerably different than it is today. At that time, research was primarily conducted at single institutions and proposed by single investigators. Quite a different pattern of research emerged in the 1980s and 1990s, involving large, multicenter trials whose protocols are developed and supported by central sources. This creates a great deal of redundancy in the system of institutional review board review.

While such redundancy has both pros and cons, consider the case of a multicenter trial held at 20 different institutions. It would be very inefficient to have 20 different institutional review boards attempting to enter data into a system. I think it is more reasonable to expect a data coordinating center or a funding source to accomplish that kind of data entry.

Second, for ongoing trials, the timeliness of data entry is a consideration. By regulation IRBs must review research at least annually, but they can do so more often if warranted for the protection of human subjects or if there are problems. This may mean, however, that the institutional review board would not be aware of precise data on numbers and characteristics of participants enrolled, or data on which the trial is closed, or other important data, as soon as it becomes available in order to include it in a registry as soon as possible. Again, the research team or data coordinating center is in a better position to do this type of data entry.

Another consideration is the choice of registry that data should be entered into. For example, should data from a trial dealing with Kaposi's Sarcoma go into PDQ? AIDSTRIALS? the NIH registry? Internet? the registry of registries that our colleagues from the University of Maryland spoke of? Theoretically, such a trial could go into all of these. Registry redundancy may be another danger facing us.

A related consideration is administrative and financial burden, especially in an era of scarce resources. If a Federal agency were to make additional regulatory reporting requirements upon institutions or investigators, the Office of Management and Budget may become involved and require compelling justification on the basis of cost benefit.

Finally, I wish to speak briefly in general support of a statement that Dr. Moher made yesterday concerning dissemination of research data. He suggested that it is unethical to conduct research and then not publish or otherwise disseminate the results. Unpublished research investigations use scarce resources and expose participants to risks and inconveniences without providing the benefit of knowledge to the research community or public. The institutional review board has had only a limited

role in clearing data to be published and in requiring dissemination of trial results in registries or elsewhere. IRBs might, however, be hesitant to approve further research by investigators with poor track records in following through on publications of studies.

As suggested yesterday, there must be a process—perhaps a journal peer review process or an NIH clinical alert assessment process—that provides a check on the integrity and quality of the results of completed trials before they are entered into the registries. I would not propose that unanalyzed results be reported through a registry. There are implications for disservice to the public and to the research community if the data to be entered are not carefully screened for quality.

REGISTRY MAINTENANCE AND QUALITY CONTROL IV

Peter Fayers

British Medical Research Council (MRC)

DR. FAYERS: I am going to talk about the UKCCR (United Kingdom Coordinating Committee for Cancer Research) registry of cancer trials and the quality control and maintenance of this registry.

The UKCCCR was established some years ago to address the problem of the lack of coordination among the many organizations in the UK conducting cancer trials. It is an umbrella organization largely funded by three prominent agencies in the field: the Cancer Research Campaign and the Imperial Cancer Research Fund (both charities), and the Medical Research Council (which is government funded).

One of the first actions of the UKCCCR was to establish a trials registry in 1985. At that time, computers were not as widely used in the UK as they are today, and the registry took the form of a large, telephone-directory-style book. Because of all the work involved, it was never maintained and eventually died.

In 1991, the UKCCR provided funding to the MRC to produce a new registry in the form of an online database which would be maintained. Currrently, the registry contains 130 open trials and 400 closed trials. We know this is only the "tip of the iceberg" in the UK, and we are trying to collect as many ongoing and closed trials as possible.

The emphasis of our registry is to collect all randomized trials of cancer therapy. The registry is trial-based, rather than publication-based, although we collect details of publications, as well. We are primarily interested in open trials, which can be either Phase III or Phase II, but must be randomized. We collect information about closed and completed trials, as well as all the trials ever done or being done in the UK. We try to locate publications for completed trials, and we update the database as publications arise.

The main difference between our registry and PDQ, for example, is that we try to be comprehensive. We make no judgement about trials, nor do we try to interpret them. If someone characterizes his or her trial as randomized, we include that in the registry. In contrast, PDQ actually reviews the quality of trials and gives approval. It is considered a form of recommendation for a trial to be included in PDQ. We have access to PDQ in various ways, and these systems are complementary in many respects.

I would like to emphasize that ours is a trial-based registry, and the main entity we are trying to catalogue is the trial. If any publications result from a trial, we will include them, but we are mainly concerned with indexing other things by the trial.

As I said, this registry was established because people were confused about studies being done in the UK. This situation has improved greatly. My concern now is what is happening in the rest of the world. I found it quite interesting yesterday when somebod mentioned a large prostatectomy study, because I wonder if that person is aware that we are launching a prostatectomy trial involving 1,800 patients in the UK very soon. The crucial problem now is finding out what other people are doing in other countries around the world.

In fact, the European Community (EC) is funding a European-wide registry which came into existence this year. They are adopting our software, so there is complete compatability between their registry and ours. The EC's intent is to provide encouragement and seed money to enable other European countries to establish registries. The goal is to develop a database that people can access for information about trials in European countries.

We are also collaborating and communicating with the Cochrane Collaboration. Ultimately, we are interested in what happens on a worldwide basis. As I work in the field of cancer, I often find it convenient to think in terms of particular disciplines. What we really want is a worldwide database of cancer trials. In that sense, we see the objective of our work with the UKCCCR registry as being our contribution to a worldwide registry.

Our registry will be made available free of charge as an online document on a UK academic network called JANET. There is a network of cancer data centers in Europe known as EuroCODE which would be a natural medium to distribute our registry, as well as the above-mentioned European registry. In each country, there is to be a central computer that will allow clinicians to learn about trials and to enter patients into trials. We will also be available on Internet, via dial-up modem lines, and in printed form.

I will turn now to how we address registry maintenance and quality control. The actual meaning of quality is open to discussion. To me, it means that the registry suits its purpose and that users rely upon and trust it. I see quality as consisting of five elements: completeness, no duplication of entries, accuracy of entries, currentness, and trials of "good quality". I do not think these five aspects are equally important, and consider the most important to be the completeness of the registry. Our purpose is to let people know that certain trials exist in a particular disease area. If a user wants to find out more about a trial, it is up to him or her to contact the investigators, collect and read the publications, and so forth.

The problem is that, in most countries except Spain, trials registration is not legally required. Clinicians consider the process of registration difficult and tend to avoid it, making the task of those trying to establish a registry that much harder.

In response, we have tried to make the registration process as easy as possible, and are seeking only a minimal dataset. We consider the core content as discussed yesterday to be somewhat lengthy and have eliminated some of those items in order to simplify the registration process. In this way, we hope to achieve as complete a registry as possible. Anyone interested in further information about the trials can obtain that later.

We also do not criticize or make judgments about the entries, as I mentioned earlier. If someone characterizes a trial as randomized, we accept that without further investigation.

Currently we are targeting the major trials offices, in order to establish our credibility. We will then contact smaller groups to solicit their participation. If necessary, we will publicize the fact that certain groups are not cooperating in hopes of achieving the most complete registry possible.

Unfortunately, it is still very difficult to ensure completeness. We will also be collaborating with various interest groups in the UK that are familiar with trials being conducted. Other sources we will use are ethics committees, the National Health Service R&D registry, college departments and similar groups, literature searches that allow us to work backwards from published trials, and various meta-analyses and overviews. Ultimately, we will need a lot of publicity, hard work, and constant nagging of people.

Another aspect of quality in a registry is avoiding duplication. We would not want to mislead people and tell them about two different trials which, in fact, are one and the same. This, perhaps, is not a tremendous problem now, but it is worth bearing in mind as registries are developed. In fact, I think registries help to reduce this type of problem.

I will conclude by expanding upon our practice of not interpreting or judging the quality of entries, which differs from some other registries. Our major concern is to highlight the existence of trials and to develop a comprehensive trials registry. We do provide information about the nature of the randomization that will help a perceptive reader decide whether or not it has been done properly. We provide that information, but we do not interpret it. Evaluating the comprehensiveness of a registry is a continual process which must be addressed in a variety of ways.

REGISTRY MAINTENANCE AND QUALITY CONTROL V

Jean-Pierre Boissel, M.D.
University of Lyon

DR. BOISSEL: The maintenance of the International Society of Thrombosis and Haemostasis (ISTH) registry is based on three items. First, we relate work of investigators and sponsors. Second, we conduct a yearly mail survey. Third, we request that the investigators and sponsors in our network complete a questionnaire. (They can also send a protocol of their study if they wish, but this is optional.)

Each completed questionnaire is manually checked for completeness and consistency. We do not check for duplication, because we think our main goal is to register every ongoing or planned trial. We often have to call for additional data on incomplete items and incorrect answers, especially regarding randomization.

Once the questionnaire is considered complete and correct, it is entered into the computer. An annual report is prepared in a tabulated format and checked by two independent reviewers. We compare the tables and the original questionnaires.

I will now discuss some key issues surrounding dissemination in registries of planned and ongoing clinical trials. One is that access should be available to all who want to use the data, including consumers, physicians, and researchers. Second, accessing the registry should be easy and require no special skills. If possible, it should be guided. Finally, each access should be recorded as an aid to assessing the registry's utility.

The main method of dissemination for the ISTH registry is an annual report which used to be published in Thrombosis and Haemostasis and is now published in Clinical Trial Meta-Analysis. We also answer queries by telephone, fax, or mail.

The advantages of an annual report include that it is very inexpensive and readily available. There are, however, strong disadvantages. One is that the updating is only periodic. Also, to get the report, one subscribes to a journal rather than the registry. Thus, there is no direct contact between the users and the registry keepers. This makes is impossible to assess the registry's frequency of use.

This led us to consider telematic access and, for the past year, we have been using the Minitel system. This is a telematic network available to everyone in France who has a telephone. It allows users, at no extra charge, to do their banking, order event tickets, book airline flights, and more from the convenience of their homes. In addition, they can access the ISTH registry of clinical trials, provided they have the keyword. Perhaps the biggest disadvantage of the Minitel system it is that it is not well-known abroad, despite the fact that it can be accessed from every country.

Updating by computer network is an attractive alternative with strong advantages. It would permit continuous updating of the registry, as well as online quality control. Completing and verifying questionnaire responses would be much easier with such a system, also. However, there is at least one significant disadvantage, and that is that it provides no automatic reminder as do the yearly surveys and accompanying questionnaires.

Currently we have limited access through the VIDEOTEX, and the X25 can be used from every country. Very recently, we added access through Internet to ease the access from foreign countries.

As mentioned, the ISTH registry relies on manual checking of questionnaires for quality control. The most frequent errors are empty boxes. Other frequent problems are a complex description of the study treatment, confusion about factorial design, and errors surrounding randomization procedures.

I will conclude by raising the issue of how to assess registries of planned and ongoing clinical trials. Such an assessment should include an estimate of the registry's sensitivity, meaning the number of false positives and false negatives (missed trials). I believe that the number of false negatives is more important, although it is very difficult to evaluate. Last but not least, we should evaluate the utility of the registry by assessing the number of satisfied queries. This, too, is quite difficult to accomplish, but may be made easier by telematic access.

DISCUSSION

DR. I. CHALMERS: Dr. Porter's presentation brought to mind a lesson I learned from Curt Meinert about 5 years ago. He noted that using the word "subject" to describe people participating in controlled trials demeaned those people and suggested they were being subjected to things. they wish to be participants in this process. Four years ago, the British Psychological Association decided to stop using the word "subject" to denote participants in controlled trials. I agree that the word should be dropped, because it does not reflect the sort of atmosphere that all of us would wish to encourage.

DR. PORTER: I agree with you. In fact, we had a discussion about this recently in a Federal agency coordinating committee meeting. I believe the word "subject" comes from the same root as "subjugation" or "subordinate" and has the underlying effect of indicating that people are not full participants in the research in which they are involved. I also do not like the terminology "use of human subjects," "human use committees," or the word "victim," among others. I agree that it would be helpful to avoid the use of "human subject."

However, having said that, the word does permeate our regulations, law, and historical documents, so it can be difficult to avoid. Sometimes the words "participant," "volunteer," or just the word "human" can substitute for "human subjects." I did note that, in quoting directly from the regulations in my presentation, I used the term "human subject" quite a bit, and I thank you for that comment.

DR. GREENHOUSE: My question concerns how confidentiality and ind ividual privacy are protected in registries. Are names and/or identity numbers entered? How does one balance the privacy of an individual against dissemination of data to anyone using the registry?

DR. BOISSEL: In the sort of registries we discussed this morning, there is no data regarding individual participants. All the data concern general characteristics of trials, not individual patients.

DR. LEVINE: I am Jerome Levine from the University of Maryland, Department of Psychiatry. During the 1970s and half of the 1980s, I was in charge of the National Institute of Mental Health's Extramural Psychopharmacology Research Program, which supported clinical trials on treating psychiatric disorders with medications at centers around the country. We created what, according to Curt Meinert's textbook, was one of the first registries of clinical trials. We had a research plan report (in essence a protocol of a trial) which was sent centrally to the National Institute of Mental Health, first by grantees of the Institute, and then by whoever wanted to have data analyzed by our biometric laboratory.

Over a period of some 10 years we created a registry of about 1,200 trials. The data did not describe individual patients, b was what one would expect in the plan of a trial. Later, we also created a research completion report that described trial results.

Today's computer technology makes it much easier to put together a trials registry now than it was earlier. The critical question has become, what is the use of the registry? One must determine who the users of the registry will be and what questions they will ask of it. The remaining details, including how and where the information is to be collected, will then fall into place.

At th ne, our pyschopharmacology registry served as a useful entity for the research community. People referred to us as the "Yellow Pages of pyschopharmacology" and would call us to learn what clinical trials were ongoing. We were far from complete, but we probably had the best picture of what was going on. This made us useful to our constituents, and they were willing to make the information available to us.

I do not know if anyone has addressed the question of obtaining information about trials supported by industry. For example, information is generally available for trials supported by the Public Health Service or individual investigators. For trials supported by industry, however, (and that is a large number of clinical trials, especially for new drug treatments) we were never very successful in getting good reporting.

My final point is that the question of how the registry is to be organized is important. For example, it could be done by content area or interest group, such as psychopharmacology or cancer. Another option would be to use clinical trials as the rubric, either nationally or internationally. Again, it depends on the questions to be answered and the information to be provided by the registry.

DR. LARSON: I am Kristen Larson from the University of Maryland. I want to address Dr. Porter's question about with whom IRBs or investigators should register their clinical trials, because the subject area and scope of various registries differs. For example, a trial concerning Kaposi's Sarcoma might be interesting to Peter Fayers, AIDSTRIALS, and others. The Registry of Registries, however, would not be interested in it. We are interested in becoming a "Yellow Pages" of registries, not a "Yellow Pages" of clinical trials.

Something that the Registry of Registries would like to do is to facilitate communication between registries when there is a trial or series of trials that would be interesting to more than one group. Creating linkages between individual registries, as the UK Cancer Trials Office and the AIDSTRIALS offices are doing, also helps to avoid duplication.

DR. MOHER: One of the best features of the Online Journal is its electronic threading, which is another mechanism for avoiding duplication. It would not be particularly problematic for a trial to be

in two registries, as long as they were electronically threaded. Then, when the trial was accessed, the information that it was also available elsewhere would appear immediately.

DR. ORZA: One potential problem of organizing a registry by disease, body system, or the like is that the people who have been mandated by Congress to look at trials concerning women or minorities would have to review all of the trials. We may need to consider other paradigms, too.

ISSUES, CONCERNS, AND CAUTIONS RE ESTABLISHING AN NIH REGISTRY I

Robert J. Temple, M.D.
Food and Drug Administration

DR. TEMPLE: I am speaking as someone who is both responsible for regulating trials and a principal user of trial results. The discussion thus far has touched on two separate matters. One is how to use a registry of published trials to learn more about a variety of important medical questions. The second is whether a registry of ongoing and otherwise unpublished trials could be valuable, not for the information already available, but because it could lead to insights about what is coming or needed.

I will first address registries involving published studies. I have nothing to contribute on the matter of search methodology, although many others are aware of those problems. I do have some thoughts on the use of these kinds of data, some of which have already been mentioned, and trialists are certainly well aware of most of these kinds of problems.

The reason for assembling registries of published trials is to be able to examine their overall results. Whether this is done trial by trial as in the past, or by some formal pooling, it is a meta-analysis. It is intended either to prove something or to create one or more hypotheses for further testing, and certain caveats apply.

The first such caveat is "garbage in, garbage out." Having many studies does not mean that quality is irrelevant, although sometimes there is a tendency to value data for its mass alone. As in any other trial analysis, one must be concerned with the details, including how randomization was achieved, how and when the endpoints were chosen and their definitions, when and by whom endpoint assessment was done, whether it is truly intent to treat, how blinding was maintained, and the like. These are all important questions, and the literature rarely provides complete answers. One must either guess or increase the statistical demands.

What is published may not equal the actual results. The differences may be small or major. It is possible for the same kind of bias to appear in multiple studies, so even a meta-analysist with a considerable amount of data must be concerned with these types of problems.

The selection of studies to be included in meta-analyses should, ideally, be neutral. The person making the selection should not know who the study authors are and what the outcomes were. This can be difficult to accomplish and not all meta-analysts do it, but it is a concept that is worth discussing.

The report of a meta-analysis should always describe what has been included and what has been omitted, and why. Only well-designed trials should be included. This would ordinarily mean randomized trials, although can historical controls may be appropriate in some cases.

Concerns about quality increase when the statistical tests approach the conventional p = .05. Unknown studies become more significant when meta-analysis results are statistically marginal. For example, Dr. Thomas Chalmers and his co-workers recently conducted a meta-analysis of studies involving the use of quinidine to maintain normal sinus rhythm after cardioversion. None of the trials were intended as mortality trials. When pooled, however, they revealed a mortality disadvantage for quinidine, along with a considerable benefit in maintaining sinus rhythm. The numbers involved were small—12 versus 3. That is enough to reach a certain level of statistical significance, but some additional trials could help to solidify the results. The lower the p-value, the less of a concern this is. One also need not be so concerned about the multiple endpoints problem with a low p-value. Searches through already-conducted trials would likely find a variety of endpoints of interest, and a very low p-value covers a lot of multiplicity testing.

There may be an inherent bias in the conduct of meta-analyses, and I would be interested in discussing this. I believe the afficionados in the field usually know how the meta-analysis is going to come out, because they are aware of the largest trials. It also seems possible that meta-analyses tend to be conducted when people already know the answer. Once again, a lower p-value is reassuring.

Finally on this topic, meta-analysts should provide follow-up to their reports in the literature. In my opinion, the resurrection of meta-analysis from what we used to unfavorably call "unplanned pooling" is attributable to the rigorous efforts of Richard Pitot, who emphasized follow-up, the analysis of oral studies by an intent-to-treat approach, investigators' providing complete follow-up on patients who were not included, and more. It is critical to maintain the view that this is not an area for casual behavior.

I will switch now to the issue of registries of ongoing trials. This is a more formidable task because it involves creating something entirely new, rather than searching for trials that are already available.

This makes defining the goals especially important, and there are several. One is to provide access for patients and physicians with a problem for which there is no treatment. I wonder whether we are not already in a good position on this goal, because the major problems have been addressed. AIDS and cancer are probably two of the most important conditions that are not successfully treated by available therapy, and we already have the PDQ and AIDS registries. Serious neurological diseases represent another category where there is tremendous interest in any product being tested. The relevant specialty and patient societies tend to know about ongoing trials. They have their own

information networks, and I suspect that significant trials in multiple sclerosis, ALS, or Alzheimer's disease are fairly widely known soon after they begin.

A second major use is for investigators and funding agencies. Avoiding repetition is important. One usually would not want to do a study exactly like another that has been done already. It may be helpful to know there is a larger database likely to be available in the near future. We talked yesterday about whether there will be a large enough database to scrutinize subsets of gender, race, and so on.

Another major use is to enhance meta-analyses, by providing information about additional large trials that may have been unknown previously.

A use that has been important in some cases but that is not discussed often is to ask new questions that the database had not asked previously. One example is asking about subsets of the whole population, which larger databases can do. In other cases, databases have been searched for questions that were not initially posed. I think that is true of the Chalmers meta-analysis of the quinidine studies. They were not mortality studies but, when they were pooled, they provided a mortality result. For a registry to be useful, it must have both government and industry sources of trials. Further, as Dr. Boissel said, international trials must be included.

Again, we are talking mostly about randomized trials. One can create a hierarchy of the types of trials to be included in a registry. The most obvious choices would be those with an explicit mortality or irreversible morbidity endpoint. Learning about these studies early is important, because such information could influence the choice of other trials to be conducted. Some might think a registry should include only those trials, but I think it is worth going further. Indeed, any trial of substantial duration or size is worth making available—not necessarily because of the endpoints being studied, but because, when pooled with others, such studies may provide additional significant information.

In general, it is less important to include earlier, smaller, or poorer studies, except where they relate to access. Thus, we are mainly talking about trials that would be considered Phase II, III, or post-marketing controlled trials.

Trials conducted by industry are very important, although this is not always recognized. Especially in the biotech era, access to many interesting drugs comes in the form of industry-controlled trials, as is the case with drugs intended for use in neurological diseases. In addition, there is a growing number of large, "government-style" trials being carried out by the drug industry. Examples include the Gusto trial, two major trials of ticlopidine (a platelet antagonist) with some 1,000 to 2,000 patients, and post-procedure and post-thrombolytic trials that are being entirely industry run. The registry would be incomplete if such trials were to be excluded.

The access problem is real, as we learned in yesterday's discussion of the AIDS registry. Much of this information is proprietary and protected by law. On the other hand, I doubt that any large, multi-center trial that industry conducts is truly secret, and I wonder whether the idea that those are proprietary and not known to people is more myth than reality. Perhaps industry would agree to register trials that were reasonably far along, or after they were completed or submitted in a new drug application (which probably is the first time they would be useful). We would want to know about large, mortality-endpoint trials earlier, however, because their existence could affect the decision to conduct other trials.

There is significant interest in enacting legislation that would require registration of certain kinds of trials. Initially, however, it may be worth exploring voluntary listing of trials. Perhaps this is an area that we could contribute to.

AN INITIAL RESPONSE OF A PHARMACEUTICAL INDUSTRY PHYSICIAN TO THE PROSPECT OF INDUSTRY STUDIES INCLUSION IN A CLINICAL TRIAL REGISTRY II

L.D. Edwards, MBBS, FFPM, Chairman Special Population Committee (PMA) and
Assistant Vice President Clinical Research Hoffman La Roche

DR. EDWARDS: I must first state a disclaimer. I do not represent the opinions of Hoffmann La Roche, nor a definitive position of the PMA. This meeting is examining the concept, defining the need, extent, and purpose of a national registry, as Congress directed the NIH to do. Until this is complete, it is hard to formulate a response. However, in preparation, I surveyed a few fellow members of the PMA and individual members of firms in the Association for medical, scientific, marketing, and legal opinions to obtain a quick response that identified problems and potential issues which must be addressed before industry is involved.

The first reaction was that this was a hangman's noose for commercial interest and profitability. Once this reaction receded, the overall conclusion was that we are not opposed to NIH and academia formulating a Registry of Clinical Trials. Our concerns center around what industry contributes, who has access, and to what this access will lead. What is in it for us?

Government and academia have different purposes, while industry sees little benefit for its members. Government hopes to use the registries for guidance in health care reform and cost; to cover politically sensitive agendas posed by the Women Congressional Caucus, minorities, the elderly, and children's and activist pressure groups. Academia and NIH hope to use the registries to avoid duplication of grants, monitor compliance with the National Health Revitalization Act, retrieve through retrospective new data interpretation, new sub-group data information, identify areas of neglected research, and publish new data. Thus, the objective of both these bodies and industry are different. For industry, publication is indicated only to support marketing or registration needs of Latin America (which require publications). Industry's studies are by necessity and regulation, which often require two well controlled studies. Thus, these publications often do not contribute exciting data. Journal editors are more responsible for the rejection and suppression of these types of studies than "dark sinister forces" alluded to by prior speakers. Indeed, many other sinister reasons for industry's reluctance to share early plans and data have been expressed at this meeting. I hope you will excuse the pun, but, the real reasons are industry's dexterity in exercising its rights, rather than being sinister.

We need to maintain our "intellectual property rights," not just patent protection against the "product pirates" of Latin America, Pacific, and Eastern Europe, but also protection of our ideas,

applications, processes, new indications, and new directions from our US competitors. Will Academia and NIH also yield their proprietary and intellectual property rights so willingly?

We need to protect our shortened patent protection times. Long development times from patent to approval, averaging 12 years, leave little exclusivity time for profit which must be compressed into an average of 5 years. Thus, information on manufacture, analysis, specifications, research plans, competitors, dosage and clinical studies which might guide competing companies and, later, generic companies to invent a new product or replicate an end of patent product, will never voluntarily be yielded. We are in a competitive environment now sharpened by health cost reform–20,000 jobs in the industry have been shed in 1993 alone. This competitive environment is intensifying. One firm has as much as 7.5% of the US market (a clue is it starts with M; this and 19 other firms comprise 75% of the market.

Our pharmaceutical industry outspends the NIH by $12.6 billion to NIH $9.7 billion in 1993. Many of you may not be aware that for each drug approved (and only 1 of 5 in clinical study make it to full approval), $359 million dollars is spent. Part of this is money that would have been earned if the capital had been invested. These are not my figures but those of a US Office of Technology Assessment report (1990).

Of the 12 years an average drug takes to get to market, FDA review lasts about 20% of this time. Thus, the much vaunted user fee will at best save 2-1/2 years, maybe 1 to 1-1/2 years, and I predict will result in a higher rejection rate. Review of the FDA's statistics for the total number of new entities and their average review times over the years indicates that it is not surprising that firms wish to "keep their powder dry" and retain competitive advantage by not yielding data on ongoing and planned studies. With up to 6 years of preapproval, this negates commercial surprise and allows competitors to interfere with approval and launch. More of our money is spent in developing medicines than the NIH's, as the NIH concentrates on basic research. However, we need profitable compounds to drive further investment in research for newer and novel agents.

In the USA, the pharmaceutical industry actively contributes information on ongoing, completed, and published studies with the AIDS and Cancer registries. It must be understood this is made possible by reducing extensive requirements and shortening development times (averaging 12 years for other agents). This makes it harder for other companies to respond.

In conclusion, industry does not withhold data from publication because it's "against the drug"–"publication bias". It is withheld because the data may be less publishable (27 drug interaction studies demonstrating nothing makes poor reading) and is duplicated (of 2 well controlled studies, one is accepted). In addition, META analysis is regarded by the industry with suspicion, so often a

negative interpretation is recognized rather than positive findings ignored. It was refreshing to hear some positive affirmations of medications yesterday from Dr. T. Chalmers.

I believe both government and academia must recognize industry's legitimate concerns. Dr. Temple's suggestion of legislation similar to the 1989 "Hope Legislation" (forcing industry to declare ongoing studies open to patient access) would not be acceptable if broadened to more therapeutic fields. In my opinion, this would result in at least Phase I and probably much of Phase II research being moved outside the USA.

As a concerned, involved individual in the US industry, I applaud your current efforts–give me your proposals to take back to my association and their member firms. But I would urge, as did Dr. Muir Grey yesterday, that your requests be brief and simple because lists such as Lois Colainni's, that contained 87 items will never be accepted. Clearly, we can help with published studies, even with data often described in promotional material as "on file," but not hold your breath. Much of the "hidden data" will be small, non-pertinent, and regulatory required studies.

ISSUES, CONCERNS, AND CAUTIONS RE ESTABLISHING AN NIH REGISTRY III

H. Schoolman

National Library of Medicine

DR. SCHOOLMAN: Given the support for the idea of creating a registry of clinical trials, I will discuss how we might proceed toward that goal. For a variety of reasons I believe that, rather than creating a registry of clinical trials, the most practical approach is to create a network of registries of clinical trials. Throughout the world, there will be many types of registries with highly specialized purposes, connected by Internet. Thus, there must be linkages that allow the registries to communicate effectively with each other. Each registry will establish access devices to serve its own constituency. Within the network, however, there must be clear and effective linkages that allow the registries to share each other's databases.

For trials in which pharmaceuticals are the subject, for example, linkages from CIS registry numbers would be very effective. That is not adequate, however. A reliable, updatable standard nomenclature is needed—perhaps MeSH, ICD9, or a maintained international nomenclature. The notion of adding a core set of elements is appealing, but difficult to implement beyond a small number. Whether or not there is a core set of elements, each registry and database in the network is likely to contain many unique elements.

People have suggested that the workload of creating such a network of registries be divided into published trials and ongoing trials. I think the best choice for the registry of published trials is MEDLINE, which already contains a high percentage of published trials and will continue to do so. Rather than creating a new registry of published trials, time and effort would be better expended on improving access to and retrieval from MEDLINE.

MEDLINE also provides an immediate linkage to the literature surrounding trials. Also, delivery of trial reports is guaranteed with MEDLINE. Even if a trial is not published in one of the 3,700 journals that MEDLINE indexes, it is likely to be published in one of the additional 12,000 journals that MEDLINE holds.

The identification of unpublished and ongoing trials can only be guaranteed by an initial registry of trials in general, or by identifying all trials as they begin. A major problem with having a registry of trials at their commencement is how to achieve the input. We have heard why it is difficult to achieve such input from industry. While gaining input from academia and governments is perhaps a bit easier, it is not without major obstacles. I suspect that funding agencies would be more useful in assuring the input of investigators than in providing data themselves. On the other hand, funding

agencies have requirements relative to the expenditure of their funds that Congress mandates they report. Thus, there is a certain amount of information that they necessarily will provide.

In any event, both carrots and sticks must be used to evolve a reasonably effective mechanism for identifying trials at initiation. Once that is accomplished, the technical details of following the trials can be worked out. The communication mechanisms for the interchange across a network of trials already exist and will improve. We should impose as little restriction as possible on network members and maintain only enough linkages to enable effective communication between registries.

ISSUES, CONCERNS, AND CAUTIONS RE ESTABLISHING AN NIH REGISTRY IV

Judith H. LaRosa, Ph.D.
Office of Research on Women's Health, NIH

DR. LAROSA: There are a number of issues that I think need to be addressed. First, should a clinical trials registry be limited to randomized clinical trials? I think Dr. Temple spoke eloquently on this point, and I would argue that the registry should not be restricted to randomized clinical trials alone. There are other data that are important for us to know. Dr. Levine raised the question of to whom a clinical trials registry should be available and directed. We can say it is directed to whomever we decide, but I am sure, as our patient advocates indicated yesterday, that a great many individuals will use it. Randomized clinical trials alone may not provide all the information they need. In a practical sense, however, how much more we can do is another question.

This leads to my second point, which is how should such a registry be implemented, maintained, and supported? It would be marvelous to have a complete portfolio on all clinical trials that have closed and are ongoing, but we heard about the tremendous problems that would involve. We must include some basic information on each trial. Exactly what should be included is debatable, but we will fail if we must continually ask investigators for additional information. I would argue for a minimum onto which we can build a maximum later, when the registry is established and people believe in it.

It is crucially important that this minimum include some of the study questions, to enable sub-analyses. We need to know where we want to go in the future, and this relates to the question of what constitutes substantive evidence. I would also argue, as Mr. Baker suggested yesterday during his discussion of how patients might use the registry, that the information must be written in cogent, comprehensive, non-jargon English.

What about the ethical issues involved in implementation? Dr. Porter has discussed these at length. Perhaps having ongoing information in a clinical trials registry can help us to avoid problems such as the DES study done in 1977. It will also allow us to examine Dr. Greenhouse's question of whether or not pregnant women have been included in clinical trials and what has happened. (In fact, we do have pregnant women in clinical trials. The AIDS clinical trials have sought to enroll pregnant women to study vertical transmission of HIV/AIDS.)

Turning now to maintenance and support, this cannot be a one-organization effort. We must collaborate. The University of Maryland's Registry of Registries is crucially important so that, when investigators or others want to know where information is being held, they can find out easily. Dr.

Moher's information on Internet was absolutely essential. How can we get into that and extract the information that you as a patient, scientist, health care provider, health care planner, or anybody else wants to know? Congress, too, needs to see that we are collecting information, if we want to avoid legislation on this.

We would hope to find ways to include the pharmaceutical industry, as well. I agree with and understand Dr. Edwards' point of view, but we need to have some of that information. How can we make this a win-win situation for government, academia, and industry?

Finally, how should the information be accessed and disseminated? I view access as a relatively passive action. One can go and get information that is there for retrieval. That is vitally important, and Internet and other systems allow us to do that—as long as the information is available.

Disseminating the information is a more active process. What mechanisms should be used to proactively disseminate the information to those communities of interest that need to know about it? Further, who should decide what is actively disseminated?

I am asking that we all join together in this extremely important effort—first, because it is important for all of us, and second, because Congress is demanding it. We must develop a system that is workable; that will not fall apart in a few years because it was not designed in such a way that people wanted to participate and keep it going.

ISSUES, CONCERNS, AND CAUTIONS RE ESTABLISHING AN NIH REGISTRY V

Joan Porter

Office of Protection from Research Risks, NIH

DR. PORTER: As the last speaker at this meeting, I would like to summarize our discussions by asking a series of questions that must be answered before we can establish another clinical trials registry.

First, what is the definition of a clinical trial? Will it be a narrow definition (as Dr. Friedman offered yesterday), or more comprehensive? Dr. Temple suggested that limiting it to Phase III or Phase IV clinical trials may be an artificial distinction and, in many contexts, would not be very helpful. To fulfill the NIH Revitalization Act requirements, we will need more data available to provide information about how the genders, minorities, and ethnic groups may react differently in various research contexts. This will enable us to frame clinical trial questions at later stages.

Will we include completed trials, ongoing trials, or both? I assume we mean both. Will we include behavioral research as well as biomedical research? The answer probably is yes.

Who will use the trial registry? Will it be investigators, institutional review boards, physicians, patients, initial technical merit review groups—that is study sections, funding institutions—for purposes of allocating resources or designing multi-center trials? Will it be data and safety monitoring boards? We have not discussed at length whether or not the clinical trials registry would be of value to such groups. Will it be of use to industry? to Congress?

Why will these different parties be using a clinical trials registry? As Lois Ann Colaianni noted, the reasons for using the registry should dictate the registry's contents. What, then, will the registry contain?

Next, how will we ensure the quality and timeliness of data? Who will guarantee continued support of the registry? Dr. James suggested we had a clinical trials registry up and going, and then we had some support considerations.

Another question is, who will enter the data? What are the costs and benefits of establishing a registry? Are there concerns about confidentiality? about proprietary interests? What registries should be contributed to by the various parties, and which registries should they use? What authorizing mechanisms operate to make a registry complete? law? regulations? institutional policy? funding contingencies? IRB approval contingencies? good will? peer pressure? journal publication procedures?

We will need to answer all of those questions before we can begin to establish a registry.

DISCUSSION

DR. DICKERSIN: Dr. Edwards, you began by asking how the registry might benefit the pharmaceutical industry, and I am not sure I heard an answer to that question. You concluded by saying we would have to make it worthwhile to have a registry for the industry. I was just wondering what types of things might make this more appealing to industry?

DR. EDWARDS: If there were some way we could utilize the registry to save time and duplication, or could utilize some of your studies in support, then that would be very encouraging. Information also is a very valuable commodity, so access to others' information would be of benefit. However, we must protect the proprietary issues—our database, indications of the timing of when we are likely to bring a drug on the market, and the like.

DR. ORZA: I am from the GAO. I have one comment and one question. My comment is about medicine and public health. My background is in public health; not medicine. Those two perspectives are compatible, but different. We have been focusing on medical users of the registry, but we should also keep the public health users in mind—with regard to both the content and organization of the registry. The public health questions that will be asked will require the ability to access in different ways. A public health question would not necessarily be searched by disease. For example, someone might be interested in all trials examining acupuncture, community interventions, or children. We should keep those kinds of users in mind.

My question is about proprietary data and access to it. Do patients who participate in an industry-sponsored trial have a legal right to know the trial's outcome?

DR. EDWARDS: They usually do, through the investigator. I am not sure about the legal aspect, but I can find out. I would imagine so.

DR. I. CHALMERS: I am from Oxford. My question is for Dr. Porter and concerns the role of the institutional review board (IRB). I was on an IRB for about 4 years, and I cannot remember whether there was any formal statement of who the IRB was meant to serve.

Do IRBs in this country see it as part of their duty to serve the proprietary interests of the drug industry? This is not meant facetiously, because I can envision two ways of approaching how patients are recruited to trials which may never be publicized. One is to insist that those trials be registered in some way, so people know they exist. The other is to insist that patients invited to participate in such trials be told there is no guarantee the information derived from their participation will be made public. I see no other option for IRBs, if their main responsibility is to participants.

DR. PORTER: I would say yes, the major responsibility of the institutional review board is to protect human subjects in research. I know of no instances in which participants were informed that

trial information would not be released to the public. That issue is not well thought through or added to the informed consent documents. IRBs are mandated to scrutinize the risks and benefits to individual subjects, and the benefits to the development of new knowledge. In weighing the risks and benefits, they should factor in the extent to which data generated from a study will be available in the public domain.

DR. TEMPLE: As Dr. Edwards said, not all trials that industry conducts are of ground-breaking significance or even very interesting. For example, a trial that proves that another ACE-inhibitor lowers blood pressure may be of interest only to the company seeking to register it. It would be hard to get such a study published unless there was something unusual about it. No one knows what patients who participate in such trials expect about publication. Some people like to be in trials because of the quality of care they get, and the like.

I have two questions for Dr. Edwards. Under current rules a company cannot market another company's drug during the period of exclusivity that follows approval, unless it has conducted the studies itself. Marketing under a so-called "paper application" is barred by the Waxman-Hatch provisions. An interval of 5 years must pass before the company can market with no data or paper data. That, then, may not be a major impediment.

Also, the new designs and nature of the work-up of the drug are revealed in detail to advisory committees and in reviews carried out at time of approval. Perhaps at that point, which is not far from when the data actually become available, the disincentives for making it available are decreased considerably. If that were true, the availability of those data in some formal way might still be of value at that point.

Further, the trials people want to know the results of are mortality trials and the like, and they are interested in the analysis. Many non-mortality industry trials would be of interest only if it were possible to do other analyses of the database. Then the question becomes, where will the data be provided in sufficient detail to allow that? The tapes must be available.

Just knowing that a trial exists may not be enough to enable meta-analyses. The mortality results are often what is needed, and these are not always available in the trial as published.

DR. EDWARDS: The registry, then, would be a way of identifying studies that might include such data. You are correct in saying one would have to obtain the detailed data from the individual companies. Normally we eliminate patient identification other than a number which is hidden by a code.

As to the Waxman-Hatch issue and the exclusivity period following approval, that is true. However, depending on the drug, every month or so a company can go without competition is worthwhile. This has often been used as an incentive for orphan drugs or extensions.

With regard to your comment on the advisory panels, often the agency will refer a drug for confirmation of its opinion to an advisory panel, or even to get advice on how to treat the drug. At that point, the data on which the decision will be based is revealed to the advisory committee, usually in a detailed pre-report format. Some of this is usually discussed by the firm and the agency in a public forum, often at FDA offices.

Thus, it is true that studies judged to adequately support efficacy and safety are deliberated and subject to public knowledge. Many of those will be published, too.

DR. FERGUSON: So that registration of trials at that point would not be a problem for industry?

DR. EDWARDS: It would depend on the individual circumstances. Clearly, however, anything that is in the public domain could be conveyed in that format to the registry.

DR. T. CHALMERS: My concern is that we may have put too much emphasis on the problems and not enough on the need during our discussions. For example, Dr. Temple observed that we already have registries for important diseases such as AIDS and cancer. That fact, however, should not lead us to relax with regard to the desperate need for registration in other fields, including cardiovascular diseases. Consider, for example, that there are eight different treatments for acute myocardial infarction, all of which currently have ongoing randomized trials. Even that is dwarfed by the problem of hypertension, in which there are over 2,000 randomized controlled trials reported on the effects of various drugs on different mechanisms of the disease. A registry is needed to help us sort through all of this information, especially with regard to subgroups such as women, the elderly, and minorities, all of whom may develop hypertension by way of different mechanisms.

Also, I strongly object to the concept of including nonrandomized controlled trials. We have an enormous job to undertake, and including all trials would make it even more enormous.

We have a simple mechanism for "separating the wheat from the chaff" in that, if a patient has an equal chance of getting into each treatment by randomization—even though there are deficiencies in randomized trials—that limits the interference of biases in clinical practice destroying the data. Once we have a decent registry of both unpublished and published randomized controlled trials available, we can begin to think about what we may learn from the bad data. We must first have some way of getting at the good data, however, and MEDLINE has not proven to be sufficient for this purpose.

Finally, I would like to raise the problem of off-label prescription drug use (i.e., the use of drugs for therapies that have not been approved by the Food and Drug Administration) in the United States, which may be as high as 50 percent. We need to ascertain the existence of randomized controlled trials that might enable approval or disapproval of those uses. A registry is the only way to

accomplish this, because so many randomized trials have been conducted on the many uses of different drugs.

DR. TEMPLE: My impression, based on Dr. Schoolman's comments, is that we already put most important journals into a database, but need to be better able to search that database. How hard is that in the long run? Does that require a new registry, or just that we be better able to search an existing, well-constructed database? If the problem is to develop a better search capability, we do not need to discuss registries, but must discuss better nomenclature, better abstracts, and the like. That is quite a different challenge.

DR. SCHOOLMAN: I am not saying that MEDLINE as it currently exists is an ideal registry of clinical trials. If we are going to invest the resources necessary to build an adequate registry, however, it is more cost-effective to start with something that already does 80 percent of the job than it is to start *de novo*. It is also more effective to invest in an enterprise with existing support resources, rather than building something new.

Making MEDLINE an adequate registry of clinical trials would require a considerable amount of effort, not only by the library, but also by editors and searchers. We would also have to establish and adopt a standard nomenclature for what is meant by clinical trials. If we must make these investments, we should do so in the most cost-effective way, by improving retrieval from MEDLINE.

DR. GOODMAN: I am from Washington, D.C.. One speaker suggested it is impractical to consider having a registry of clinical trials. To me that equates to having said in the early 1960s that it is impossible to have a well-indexed, computerized, bibliographic database covering thousands of journals from around the world. We have seen how ridiculous such a statement would have been. Despite the many practical and technical problems to be faced in developing a registry, saying it is impossible to do so ignores current trends, as follows.

First, throughout the field of technology assessment and health services research, there is an increasing appreciation for the need for scientific rigor in analyzing clinical studies. Next, Federal efforts to conduct synthetic research (such as the patient outcome research teams that involve synthesizing all available research) are also increasing. The expertise for this exists, and the demand is increasing.

Further, we are moving to base practice guidelines that are generated or supported by the Federal government on the best available evidence. I do not know how this can be done without a clinical trials registry. Researchers such as Dr. Chalmers and others have demonstrated that lives can be saved by pulling together the best available information from clinical trials. To not do that might verge on the unethical.

Other important points are that computer technology is improving extremely fast, and further expectations are growing, too. Also as Dr. LaRosa suggested, there are national and international policy-making needs to look at clinical trials across different treatments and different types of patients to better understand how to devote national resources to health care. These overriding trends dwarf the technical problems and pessimism we have been hearing about.

DR. SIMES: I want to return to the issue of a registry of pharmaceutical trials. This workshop has been addressing an evidence-based health care system. The primary purpose of a trials registry is to provide access to all relevant trials which can help us to develop evidence-based guidelines.

We have heard that the pharmaceutical industry spends more money on research than does the NIH. Publication rights are lower amongst the pharmaceutical than NIH studies, and at least one study suggests publication bias is greater, as well. If we are serious about developing an evidence-based health care system that uses registries of trials, then we must develop a system that includes pharmaceutical trials, too.

I have two suggestions toward this end. First, perhaps we could focus on randomized trials with clinical outcomes and ignore some of the studies of pharmacokinetics and Phase I-type studies, which would swamp the system. Second, for trials where confidentiality is an issue, perhaps they could be registered in camera. Then, at some later time, researchers could locate those registered studies and learn the sizes of trials addressed to a particular question.

DR. EDWARDS: I remind you that many industry studies are relatively small, which can create problems for meta-analysis. The p-values may be nice, but the numbers of patients are small. Yes, studies are eventually registered in most countries and they do come off patent. Clearly they should be available, but that may mean waiting for long periods of time, as in the case of AIDS drugs. Was your point that, once they were registered, that data should be available?

DR. SIMES: I would like the studies to be prospectively registered before the results are known, even if we do not know at the time that they are registered, so we will eventually know the number of studies done.

DR. TEMPLE: It is worth considering which studies are most important. Those with mortality or irreversible morbidity outcomes are the most critical for our needs. These, however, are almost always multi-center trials that are widely known. (One cannot mount a 1,000 patient trial in secret.) The idea that those need to be kept confidential, then, is a bit odd.

The more mundane trials are often the ones that are not widely known. These have potential for asking new questions, but gaining access to the data can be a problem.

DR. HAIGLER: I am from NIMH. There is a publication called *Pharma Projects* that provides seemingly confidential information about ongoing trials. Perhaps this could be another source, since it provides published information in database form.

DR. DICKERSIN: First, I want to express my concern about labeling trials as mundane and non-mundane with regard to their outcomes, and that mortality or major morbidities are the only outcomes that are considered important enough to include in a trials registry. For example, in my opinion any trial concerned with stopping smoking is major in terms of mortality in this country, although the endpoint might be maintaining cessation. I urge that we think about a trials registry in general, and not by whether studies ask a mundane or non-mundane question.

Second, Dr. Edwards, you said you would feel comfortable if a pharmaceutical trial were registered once it becomes public information. Would not any study funded by a pharmaceutical company that is routed through an IRB be public information?

DR. EDWARDS: Yes. But I don't think the details of that trial...

DR. DICKERSIN: Not the details, but perhaps such information as the title of the trial, who is funding it, and the major disease being studied.

DR. EDWARDS: Yes, that is in the public domain when the industry works with the NIH and affiliated academic bodies. However, many of our investigations are conducted with non-academia.

DR. HUTH: I have two questions for Dr. Schoolman. First, with regard to your comments about the need to standardize nomenclature for international communication, what is the status of the Unified Medical Language Project? Second, I agree with the concept of improving MEDLINE if possible, rather than inventing a whole new system. What is the potential for the Library to do what it did with GRATEFULMED and create easy access to MEDLINE via Internet?

DR. SCHOOLMAN: To answer your second question first, the Library is creating mechanisms for Internet usage via GRATEFULMED. Currently, this may be an elitist approach to entering databases, rather than the "common man" approach. Perhaps this situation will be different 5 years from now. We are involved in many outreach activities aimed at easing access to the Library's databases. However, I see no reason, if the database of a clinical trials registry is created within MEDLINE, that its use could not be sequestered from MEDLINE, if that were desired. My point is that we have been engaged in creating bibliographic control over published literature for 150 years, and we are the world's primary indexers of medical literature. It seems ridiculous to ignore that and create something entirely new now.

Having said that and having reviewed steps that might improve retrieval to a satisfactory degree, I warn that it will never be 100 percent effective. Retrieval from any information system has entropy.

The only way to achieve 100 percent retrieval from any database is to retrieve 100 percent of the database, which will not do us any good.

Regarding the Unified Medical Language System (UMLS), if the registries that are to be part of this network will adopt one of the standard nomenclatures covered by UMLS (ICD9, MeSH, SNOWMED, and reclassification for our English colleagues and others), we will have a mechanism that allows us to map terms used in one database to those in another database. This will permit logical and semantic communication between databases.

In no way is this an adequate substitute for a core number of linkages that are unique and require no mapping, which are fundamental to the interchange of information among databases. Such linkages should include a unique identifier for each clinical trial in every database. However, the core interpretation can be facilitated by mapping through the Unified Medical Language System meta-thesaurus, as well as its other tools.

DR. COLAIANNI: I want to add one point. We are in the process, through the Health Services Research Group, of adding screens to GRATEFULMED that will help those interested in health services research. There is no reason why we could not develop a screen in GRATEFULMED to help those interested in searching for clinical trials, as well.

If people can authoritatively identify the published randomized control trials that are reported in MEDLINE, we will tag them accordingly. Further, if there are important clinical trials published in journals that we have not indexed, we will put those into a separate database where they could be searched. MEDLINE would alert users to their existence.

DR. HUTH: I have one more comment regarding Internet. There probably are many in this audience who use it quite readily. However, it remains grossly intimidating for many potential users. Something radically simpler is needed. We have heard about registries in Europe and elsewhere. The question is, how can we facilitate access to these for users with less sophisticated computer skills and equipment?

DR. SCHOOLMAN: I agree, and my advocacy of Internet was as a communication link between the registries involved in the network we will create. For them, Internet is an effective communication device. Each individual registry, in the fulfillment of its own unique mission, must develop access methods that are more user-friendly and more widely available than Internet currently is. Then, in response to a query to them from this more widely available access device, they can use Internet to fulfill the query by searching through the other registries on the network.

DR. GRAY: I am from Oxford. We have been discussing two types of registries—a registry of trials and registries of reports of trials. I want to speak about registries of reports of trials. First, all that I have heard in the last 2 days has convinced me of the importance of developing a

comprehensive registry of reports of trials and reviews. We will certainly be putting resources into that. We would see MEDLINE as the base, and the NLM as the focus of that activity. Speaking for my own and perhaps another region of the national R&D program, we see this as the way in which to link the offer of updating the trials.

However, there will be a need for what we might call "knowledge archaeology," or hand searching the old journals. Volunteers can search 1,000 pages an hour. This is not an impossible task, but it is a big task, and must be done.

We will be working in large chunks—10,000 pain trials, 10,000 psychiatry trials—but we would like to link with you. We also are strongly committed to the idea of a registry of reports of trials, with a separate debate on a registry of trials. I hope that this is the start of building an international registry of reports of trials and reviews.

DR. T. CHALMERS: I would like to be presumptuous enough to propose two items which might be considered the sense of the meeting. They both refer to the Public Health Service and its responsibility for the health of the people. One is that the National Library of Medicine be urged to establish easier methods for obtaining information on published clinical trials, through GRATEFULMED or whatever system they select.

The second is that the Public Health Service be urged to take measures to ensure that institutional review boards, which do respond to the requirement that they protect individuals involved in research, also respond to the requirement to report all approved clinical trials to a central agency. I recognize that money will have to be appropriated, both for the individual IRBs to process the paperwork, and for the central registry. It seems to me we must move in those two directions if we are going to have any output from this workshop.

DR. I. CHALMERS: I have a practical question. I said at the end of my presentation that every country can and should contribute to this, but that we are looking to the United States for leadership. It would be helpful if we had some guidance as to how we can help you achieve the objectives we just endorsed. For example, if journal editors knew that the National Library of Medicine was going to retrotag studies that had been published with the designation, "controlled clinical trials," they could begin looking through their journals to determine which studies met that criteria. I am asking the National Library of Medicine to tell us how we in other parts of the world can play our part in a coherent plan.

DR. TEMPLE: I would like to interject a note of caution with regard to something that was said concerning the confidentiality of IRB reviews. It is not at all clear that they can report to outside parties on what is being studied. One hundred percent of all studies go before IRBs before they are carried out, yet the information under an IND is considered confidential. Therefore, I question how

much information they could provide, or whether they could be the basis for a report. I am not saying they definitely could not be, but it is unclear.

DR. PORTER: I just wanted to comment that it depends in part on where the institutional review board is located. In some states there are sunshine laws, which require that certain deliberations be open. And to some extent, IRBs are open to the public. Now, how much data are actually shared that would be of a proprietary interest I guess are questionable. But it is not a simple picture of open or closed, available to the public or not available to the public.

DR. GRANT: I am Adrian Grant from Oxford in England. I would like to explore the issue of registries of ongoing trials, and in particular the relationship between the proposed NIH initiative and the specialist registries that we have heard about. My suggestion is that you may choose to concentrate on identifying ongoing trials and possibly leave the other activities to other players.

My comments come from a background of involvement in a registry of ongoing trials in perinatal medicine. I recognize that this process is quite cumbersome, but that the underlying problem is one of identification. Stated frankly, voluntary registration does not work satisfactorily. My view coming into this meeting was that we should try to develop generic databases of all randomized control trials. I still feel we should do that, but this meeting has changed my view in one respect. I think now that the task will require a very simple system, and I am concerned that the plans are too ambitious. Perhaps NIH should plan to develop a generic system merely to identify ongoing trials with key descriptors. Then the specialist groups could use that as a skeleton on which to base more detailed registries that would meet particular requirements for particular groups.

DR. MOHER: I am from Ottawa. One of the impressive things that I have heard today is that so many different groups are interested in trial registration. However, we should focus on being prudent and, given our fiscal resources, what we can get away with. The core content is crucially important for all users. A group such as the international collaboration of clinical trial registries, for example, could suggest a core content of 10 items. Other users could then employ a modular approach and add other information that they want.

I would select randomized controlled trials first, because they are at the top of the pyramid in terms of what we take for evidence, although there may be other useful designs. If we do have core items, we must clearly define each one, especially if this is to be an international registry that reaches across various cultures and languages.

MR. GUIMENTO: I am from the American Foundation for AIDS Research. Regarding Dr. Edwards' concerns about whether a pharmaceutical company would find it worthwhile to release proprietary information, in the case of AIDS right now, we receive Phase I information immediately from the pharmaceutical company to publish. They find that releasing the information is helpful for

recruitment, and to let people know the status of the field. Many of the summary questions and concerns being raised here have been discussed for years among AIDS treatment information providers and registries. We are now revisiting these issues as we create the international AIDS clinical trial registry, INTERACT, by combining the AmFAR, NLM, and Canadian HIV trials networks. I suggest that you review the consensus points that have resulted from the deliberations of people with AIDS and those organizations, because many of the issues are the same.

DR. WELLS: I am with the National Institute of Dental Research. I would like to return to the issue of who will use the registry and what they will be looking for, especially with regard to patient use of the registry. Let me describe the situation that we find ourselves in at the Dental Institute. Recently, some misinformation has been widely distributed via videotape, television, and print. It indicates that the U.S. Federal government funds clinical research on various conditions and problems through grants, and that patients can demand free medical care from those conducting such research. Our information offices have been inundated with people demanding information about where they can get free dentures, implants, dental care for periodontitis, and the like.

I would raise a note of caution, then. When disseminating accessibility information to patients, we must be careful to inform them that there are limits and that this is not free care. Mr. Baker with the National Association of People with AIDS yesterday noted items his group would want listed in the registry, including accessibility, what would be required of the patient, what compensation would be provided, and what other medical care would be provided. Providing information about patient accessibility will require us to expand the data items to be included in the registry.

DR. SCHULZ: I am from the Centers for Disease Control. I would like to endorse the concept of focusing on randomized control trials. Given unlimited resources it would be nice to know all the information, but I estimate that we would gather some 98-99 percent of the scientifically valid, unbiased information available by focusing solely on randomized trials.

Second, I am confused as to whether there has been a decision by NIH to go forward with a registry of ongoing trials.

DR. FERGUSON: I think the answer is no, but we have a mandate to do so for women's health. The question is how, what, and how far beyond that should we go?

DR. DUBEY: I am Satya Dubey from the Food and Drug Administration. It appears, since these registries are to be used by non-scientists along with scientists and researchers, there is a definite need to include a grading of the quality, level, and strength of evidence and recommendations. This should be done in such a way that lay people can clearly understand what kind of information has been generated from clinical trials. Perhaps the color-coding of information would enable non-scientists to give proper weight to information they receive.

DR. FERGUSON: I think we all have to keep in mind who we are doing this for. As we heard from the patient advocates yesterday, this is for the health of the public, for patients and potential patients in prevention research.

DR. TEMPLE: I think color-coding would help scientists, as well.

DR. ORZA: I wanted to return to Dr. Dickersin's comments about the idea that some trials are more mundane than others, and some are less important than others. I, too, find that concept offensive. By definition, if we are conducting a trial it ought to be important. We should not conduct unimportant trials. What are the IRBs doing if they approve trials that are not needed, or trials whose results no one cares about?

DR. FERGUSON: Perhaps it is importance to whom.

MS. CANNON: I am the manager of a contract that updates the PDQ. I have been trying to link published results with the clinical trials in PDQ, and I urge writers of these articles to include the ID number of their protocol. I often find an article that I think may be the same clinical trial, but it is difficult to link the unpublished trials with their published results without the institution's ID number.

DR. EDWARDS: I would like to respond to the "mundane" issue. Mundane in this case does not mean that the study is unimportant. It depends on important to whom, and to what audience one is speaking. It is vital that we conduct drug interaction studies, but I doubt that they will prove to be gripping reading to many in this audience. Thus, we do not mean that studies labeled "mundane" are not important in terms of developing a drug. They are not, however, necessarily important in terms of new indications, mortality outcomes, and such. I hope that clarifies it.

DR. ALLMAN: I am from Johns Hopkins. I have been working on the pain guidelines. We have been using a best-evidence emphasis, in which we ideally examine only randomized trials. If we have enough randomized trials to perform a meta-analysis, we do so. In many instances for treatments other than drugs, however, randomized trials of studies are not available. I would like to see studies beyond the randomized trial included, particularly for disciplines such as public health, and for the younger disciplines that may not have as many randomized trials for their treatments, such as physical therapy. We are trying to use an inter-disciplinary approach, and there will be people in fields other than medicine using these registries.

DR. GRAVES: I represent the Sigma Theta Tau International Library of Nursing. We have been developing a research registry for several years. We currently have 3,500 volunteer registrants, and we register both the person and the general research they do, whether or not it involves a clinical trial. We would be delighted if some kind of vocabulary resulted from this meeting. If that were developed with good definition, we could simply collapse it into our own work and then contribute.

Our goal is, within [illegible] years, to represent 90-95 percent of all research done [illegible]rsing. Thus, your clinical trials register would not meet the needs of nursing, but we can meet yours.

DR. GUYATT: I am from McMaster. We should remember (and the patient advocates have stressed this) that the prospective registration with minimal information of ongoing trials, and the subsequent availability of completed trials, is very much in the public interest. What we do with that is also in the public interest. Any way of avoiding or obstructing that is contrary to the public interest. Ultimately, the public interest must be paramount above proprietary interests.

DR. TEMPLE: I thought I heard the idea that irregular kinds of treatments—that is, other than drugs—deserve to be evaluated with a lower standard. I think the opposite is true. It is ironic that we subject things that we know work to rigorous studies, and things that lack any theoretical or other basis often undergo other kinds of (less rigorous) studies.